

BINDINGS IN WORKING MEMORY: FOCUSING, MAINTAINING, AND RETRIEVING INFORMATION

Thesis
presented to the Faculty of Arts
of
the University of Zurich
for the degree of Doctor of Philosophy
by

Laura Hein

of
Germany

Accepted in the fall semester 2012 on the recommendation of

Prof. Dr. Klaus Oberauer and PD Dr. Ute Leonard

Zentralstelle der Studentenschaft

Zürich, 2012

Pre Matúša
za držanie mojej ruky

Abstract

In this thesis I compared the suitability of different conceptualizations of working memory (WM), namely the slot (Zhang & Luck, 2008, 2009), the resource (Bays & Husain, 2008) and the three-embedded-components model (Oberauer, 2002, 2009) to account for the effects of focusing and de-focusing information in WM and the spatial distribution of errors. The results showed that non-focused information (i.e., information that has not been focused after encoding) remained in WM at baseline strength, and that de-focused information (i.e., information that has been focused recently, but is currently not focused) remained strengthened together with its bindings to its context even after de-focusing, rendering it highly accessible and recognizable. Regarding the distribution of errors, the results revealed a spatial transposition gradient, analogous to the transposition gradient in serial recall: Items spatially closer to each other are more likely to be confused than those further apart. This finding can be explained by the assumption of an overlap between retrieval cues, which cause the establishment of misleading bindings between these cues and memory contents as a function of spatial proximity. None of the three models could account for all findings. However, the three-embedded-components model gave the best fit due to incorporating the two important characteristics, which emerged from the studies: the importance of bindings in WM to account for fairly diverse effects (the fate of focused, non-focused, and de-focused information, and the spatial distribution of errors) and the importance of separate mechanisms for focusing and maintaining information in WM (to account for focusing benefits while non-focused information can be maintained without impairment).

Zusammenfassung

In dieser Dissertation wurden verschiedene Konzeptualisierungen des Arbeitsgedächtnisses (AG), nämlich das Slot-Modell (Zhang & Luck, 2008, 2009), das Resource-Modell (Bays & Husain, 2008) und das Three-embedded-components-Modell (Oberauer, 2002, 2009) miteinander verglichen. Zu diesem Zweck wurde untersucht, welche Effekte das Fokussieren der Aufmerksamkeit auf einen AG-Inhalt auf andere, nicht fokussierte AG-Inhalte hat und wie Fehler räumlich im AG verteilt sind. Die Ergebnisse zeigten, dass non-fokussierte Information (nach dem Enkodieren nicht mehr fokussierte Information) im AG unvermindert erhalten blieben und dass de-fokussierte Information (kürzlich fokussierte, aktuell aber nicht mehr fokussierte Information) mit gestärkten Bindungen zu ihrem Kontext im AG blieben, so dass sie in erhöhtem Mass zugänglich und wiedererkennbar waren. Die Verteilung der Fehler im AG folgte einem räumlichen Transpositionsgradienten, analog zum seriellen Transpositionsgradienten in seriellen Wiedergabeaufgaben: Repräsentationen, die nah beieinander präsentiert wurden, wurden mit grösserer Wahrscheinlichkeit vertauscht als weiter auseinanderliegende. Dieses Muster lässt sich durch überlappende Hinweisreize erklären, die die Bildung von irreführenden Bindungen zwischen einem Hinweisreiz und einem nicht dazugehörigen AG-Inhalt als Funktion der räumlichen Distanz verursachen. Keines der drei Modelle konnte alle Ergebnisse erklären. Das Three-embedded-components-Modell konnte die Resultate jedoch am besten abbilden, da es zwei Kernmerkmale erfolgreicher AG-Konzeptualisierungen, die sich in unseren Studien abzeichneten, aufweist: die Relevanz von Bindungen im AG, die eine Vielzahl von Effekten erklären können (fokussierte, de-fokussierte und non-fokussierte Informationen und die Verteilung von Fehlern im AG und die Wichtigkeit der Annahme separater Mechanismen für das Fokussieren und das Behalten von AG-Inhalten, um Fokussierungsvorteile ohne Behaltenseinbussen nicht-fokussierter Informationen zu erklären.

Table of Contents

Part I: Synopsis	11
1 Introduction	13
1.1 Focused and Not-focused Information in WM	14
1.1.1 Theoretical Background	14
1.1.2 Summary of Study 1: Focused, Unfocused, and De-focused Information in WM	16
1.2 The Spatial Organization of WM Representations	18
1.2.1 Theoretical Background	18
1.2.2 Summary of Study 2: Spatial Transposition Gradients in Visual Working Memory	20
2 Discussion.....	23
2.1 The Explanatory Power of the Three WM Models.....	23
2.1.1 The Fate of Non-focused and De-focused Information in WM	23
2.1.2 The Spatial Distribution of Transposition Errors	26
2.2 Two Important Characteristics of WM	28
2.2.1 Bindings in WM	28
2.2.2 Separate Mechanisms for Focusing and Maintaining WM Representations	29
2.3 Future Directions	29
2.3.1 Spatial Imprecision of Retrieval Cues.....	29
2.3.2 A Limit of Strengthened Bindings?	30
2.3.3 Computational Modeling.....	31
2.4 Conclusion	31
Part II: Empirical Studies	33
3 Focused, Unfocused, and De-focused Information in Working Memory	35
Abstract.....	36
3.1 Introduction.....	37
3.1.1 Discrete-Capacity Theories	37
3.1.2 Constant-Resource Theories	38

3.1.3 The Three-Embedded-Components Theory of WM.....	39
3.1.4 Focusing Items within Central Working Memory	41
3.1.5 Mechanisms of the Focusing Benefit	42
3.1.6 Non-Focused and De-Focused Information in WM	43
3.2 Experiment 1A and 1B: Retro-Cue Benefit and Non-Focused Information in WM.....	47
3.2.1 Method.....	50
3.2.2 Results.....	54
3.2.3 Discussion.....	58
3.3 Experiment 2: De-Focused Information in WM - Intrusion Approach.....	61
3.3.1 Method.....	64
3.3.2 Results.....	66
3.3.3 Discussion.....	70
3.4 Experiment 3: De-Focused Information in WM – Backshift Approach	73
3.4.1 Method.....	73
3.4.2 Results.....	75
3.4.3 Discussion.....	79
3.5 General Discussion.....	81
3.5.1 The Fate of Non-Focused Information in WM.....	81
3.5.2 The Fate of De-Focused Information in WM	83
3.5.3 Representations in WM are Spatially Imprecise	85
3.6 Conclusion.....	86
4 Spatial Transposition Gradients in Visual Working Memory	89
Abstract	90
4.1 Introduction	91
4.1.1 Distinctiveness in Time and Space	91
4.1.2 Ordinal or Metric Distinctiveness?	92
4.2 Experiments 1A and 1B	93
4.2.1 Method.....	94
4.2.2 Results.....	95
4.2.3 Discussion.....	98
4.3 Experiment 2	99
4.3.1 Method.....	100

4.3.2 Results	101
4.3.3 Discussion	104
4.4 General Discussion	105
5 References	109
Acknowledgements	119
Curriculum Vitae	121

Part I: Synopsis

1 Introduction

Working memory is a system for holding information available for a current cognitive task. It allows for keeping information highly accessible and for binding it into new structures, for example when mentally updating a multiple-digit number by adding another one to it. Unquestionably, WM is important for human cognitive functioning. It is involved in almost all cognitive tasks (Lovett, Reder, & Lebiere, 1999), such as text comprehension, reasoning, problem solving, decision making, and mental arithmetic. One of the core characteristics of WM is its limited capacity: Only a limited amount of information can be kept in a highly accessible state at a time. The nature of this capacity limit has been characterized differently by different theories.

Discrete-capacity theories, as the slot model (Zhang & Luck, 2008, 2009), are based on the assumption that WM is limited by a discrete amount of information that can be stored in an all-or-nothing fashion (cf. Cowan, 2005). This discrete limit is estimated to be three to four items or chunks in healthy adults (Cowan, 2001, 2010; Luck & Vogel, 1997). Constant-resource theories (Bays & Husain, 2008; cf. Just & Carpenter, 1992) propose that the capacity limit is defined by a limited resource, such as activation, which is allocated to all representations in WM to maintain them. A third approach to characterize the capacity limit is delineated in the three-embedded-components model (Oberauer, 2002, 2009). According to this model, WM capacity is assumed to be limited regarding the amount of bindings that can be maintained in WM (Oberauer, 2002, 2009), that is, the association between WM contents (e.g., a color or a word) and their contexts (e.g., spatial or serial position).

These three different explanations for the emergence of a capacity limit reflect different conceptions of WM and its underlying mechanisms; and although all three theories can account for the emergence of a capacity limit, they do not always lead to the same

predictions for other WM phenomena. Diverging predictions offer the opportunity to evaluate the adequacy of the different conceptualizations and to broaden our understanding of WM.

In this thesis, I examined the organization of WM and evaluated the appropriateness of the three models introduced above in light of the findings. For that purpose, two main questions, for which the models propose different predictions, were examined in two studies. This implies that the models' predictive or explanatory power to account for the results can be utilized to evaluate the suitability of their conceptualizations of WM. The first study addressed the question of how focusing attention on a single representation affects the fate of other not-focused information in WM. The second study addressed the question of how WM contents are organized spatially by examining the spatial distribution of errors. Besides aiding in the evaluation of the models, these questions received sparse attention in WM research so far and their answers can provide valuable insights into the organization of WM.

1.1 Focused and Not-focused Information in WM

1.1.1 Theoretical Background

In the first study, I examined what happens to not-focused information in WM, when attention is focused on another representation. Extensive empirical evidence has shown that attention can be focused on a single representation or a subset of information in WM and that this significantly improves performance on the focused information compared to conditions in which attention is not focused (e.g., Delvenne, Cleeremans, & Laloyaux, 2010; Griffin & Nobre, 2003; Kuo, Rao, Lepsien, & Nobre, 2008; Kuo, Stokes, & Nobre, 2012; Landman, Spekreijse, & Lamme, 2003; Lepsien, Griffin, Devlin, & Nobre, 2005; Makovski & Jiang, 2007, 2008; Makovski, Sussman, & Jiang, 2008; Matsukura & Hollingworth, 2011; Nobre, Griffin, & Rao, 2008; Sligte, Scholte, & Lamme, 2008; Souza, Hein, & Oberauer, 2012; Williams & Woodman, 2012). This focusing benefit is often called retro-cueing benefit, because attention is typically directed to subsets of WM by retro-cues, which are displayed in

the retention interval between memory array and probe array. All of the three models outlined above can account for the focusing benefit, though by relying on different explanations for its emergence. Depending on the conceptualization of WM underlying the explanation for the focusing benefit, different consequences of focusing attention are predicted for the remaining, not-focused information.

According to the slot model (Zhang & Luck, 2008), each object is stored in an all-or-nothing fashion in one of a limited number of slots. If the number of items to maintain exceeds the number of available slots, all items exceeding the limit are lost and participants can only guess about their identity. An extension to the original model, the slot-and-averaging model (Zhang & Luck, 2008), can account for the focusing benefit and will be referred to by the term slot model in the remainder of the thesis: Samples (i.e., copies of the representation) of the focused object are assigned to multiple slots and information from these slots is averaged at retrieval, thereby reducing variance and improving precision. However, focusing has implications for not-focused information: If more slots are devoted to keep one specific representation highly accessible, fewer slots remain available for the other memory items. This implies that focusing attention to a representation comes at a cost for the other, not-focused items.

According to the constant-resource model (Bays & Husain, 2008), a resource, such as activation, is spread across all representations, independent of the amount of these representations, to keep them in WM. The resource is assumed to be limited, so that the more representations, which have to be maintained, the less activation is available for each of them. Within this model, the focusing benefit can be explained by allocating a bigger share of the resource to the focused representation, which increases its accessibility and improves performance on this representation (Bays, Gorgoraptis, Wee, Marshall, & Husain, 2011). However, allocating more of a limited resource to the focused representation implies less available resource for the other, not-focused representations. Consequently – as the slot model

– the resource model predicts that focusing attention on a specific representation impairs the accessibility of not-focused information in WM.

In the three-embedded-components model (Oberauer, 2002, 2009), the role of bindings in WM is emphasized. The model comprises three hierarchical components: the activated part of long-term memory (peripheral component), the region of direct access (central component) and the focus of attention. The activated part of long-term memory is assumed to keep activated representations available, which might be needed in the context of a current task (e.g., keeping digits and mathematical operators available when doing mental arithmetic) and to be unlimited regarding its capacity. In contrast, the region of direct access is proposed to hold only a limited number of item-context bindings. This limit is assumed to arise due to interference between these bindings (Oberauer, 2002, 2009). According to this model, one of the representations, which are held in the region of direct access, can be retrieved into the functionally limited single-item focus of attention, where it is assumed to be highly accessible and protected from interference. The special status of this representation is reflected by the focusing benefit. Focusing attention on a single representation does not affect the maintenance of other representations in the region of direct access, because the focus of attention is considered to be a separate component from the region of direct access, not drawing from the same capacity limit. Therefore, in contrast to the resource and the slot model, no trade-off between focusing information and maintaining not-focused information is predicted.

1.1.2 Summary of Study 1: Focused, Unfocused, and De-focused Information in WM

In the first study (*Focused, Unfocused, and De-focused Information in WM*), I addressed the fate of not-focused information in WM, when attention is focused on other information; and distinguished two types of not focused information, namely non-focused (i.e., information that has not been focused after encoding) and de-focused information

(information that has recently been focused, but is currently de-focused). The three models outlined above make different predictions for the fate of non-focused information, which allows the evaluation of their predictive power to account for the findings. Little is known so far regarding the fate of de-focused information as little empirical research has addressed this question (cf. Maxcey-Richard & Hollingworth, 2012), and none of the models explicitly speaks to the fate of this kind of information.

The data of the first study suggested that non-focused representations remained accessible despite focusing attention on another representation, in accordance with the predictions of the three-embedded-components theory and in conflict with the predictions of the slot and the resource model. Non-focused information was maintained, even when it was known to be irrelevant to the task. This suggests that no focal attention is required to maintain representations in WM. Moreover, the data showed that the discriminability of non-focused items differs along their spatial positions relative to the focused representation: Confusing the focused item with a spatial neighbor was more likely than confusing it with other representations, which were spatially more distant (neighbor effect).

Although the outlined models do not specify predictions for the fate of de-focused information, three possible hypotheses can be sketched for the fate of this kind of information: its accessibility could return to baseline level, it could be inhibited, or it could remain increased (due to having been focused recently) compared to non-focused information. The data showed that the increased accessibility and recognition performance of focused representations was not restricted to the time this representation was actually focused. Indeed, the accessibility of de-focused information remained higher compared to non-focused information even after it was de-focused and attention has been focused elsewhere. This finding can be explained by assuming that focusing a representation does not (solely) increase its activation, but strengthens the respective binding between the item and its context. The

strengthened binding facilitates the representation's identification and thereby improves accessibility and performance.

In sum, in Study 1, I showed that non-focused information can be maintained in WM without focal attention, as predicted by the three-embedded-components model, and that de-focused information remained strengthened even after being de-focused. Beyond this, indicative evidence for the spatial imprecision in WM was gathered by the neighbor effect, which was then examined in more detail in the second study.

1.2 The Spatial Organization of WM Representations

1.2.1 Theoretical Background

The second question addressed in this thesis concerned the organization of representations in WM by examining the distribution of errors. The distribution of errors has been extensively investigated in serial-order research (e.g., Lewandowsky & Farrell, 2008; Morin, Brown, & Lewandowsky, 2010). In serial-order tasks, memory objects are presented sequentially and have to be recalled by their serial positions. The serial position of a representation therefore serves as retrieval cue (context cue) by which the corresponding representation from the sequence can be retrieved. A frequent type of error in these tasks are transposition errors, that is, recalling an item that has been part of the sequence, but recalling it in the wrong serial position. The transposition errors are not randomly distributed, but follow the so-called locality constraint (Murdock & vom Saal, 1967): The closer two items are presented to each other the more likely they are confused with each other.

To account for this transposition gradient, most serial recall models such as the Phonological Loop model (Burgess & Hitch, 1999), and the Serial Order in a Box model (Lewandowsky & Farrell, 2008) incorporate the assumption of limited distinctiveness on the time dimension due to overlapping retrieval cues (position markers). The degree of overlap between these context cues is determined by the (temporal) distance between items (Morin et

al., 2010). Context cues are associated to the corresponding items (e.g., linking a word to its serial position in the sequence) through bindings (Morin et al., 2010), and the more the context cues overlap, the more misleading bindings are established, which connect a context cue to contents other than the corresponding one. As a consequence, when using a retrieval cue (serial position) which is associated not only to the corresponding item, but also to other items, partial information from other items is retrieved. This increases the likelihood of committing a transposition error.

The slot, the resource and the three-embedded-components model are typically tested with simultaneously presented stimulus material, so that no serial or temporal context emerges. Instead, the spatial position of an item serves as a retrieval cue. Although the effects of serial context cue overlap are well established in serial-order literature (for a review, see Morin et al., 2010) and it seems to be self-evident to assume that other types of retrieval cues overlap in the same manner, none of the three models proposes a context-cue overlap related to the distribution of transposition errors. Quite the contrary, within the slot model (Zhang & Luck, 2008, 2009), errors are predicted to be caused by guessing about the identity of representations that could not have been stored due to the limited number of available slots. Consequently, it is proposed that errors are distributed randomly. Bays et al. (2009) reported that transposition errors, which reflect errors of mis-locations (i.e., reporting an item from the memory set, but in the wrong spatial position), occur significantly more often than predicted by chance, and make up a large amount of all errors. The constant-resource model can account for the high prevalence of these errors, but it does not predict a specific distribution among them errors. The three-embedded-components model does not specify predictions for the distribution of errors.

1.2.2 Summary of Study 2: Spatial Transposition Gradients in Visual Working Memory

In the second study (*Spatial Transposition Gradients in Visual Working Memory*), I examined whether there is a spatial transposition gradient, which reflects effects of spatial proximity between WM representations on the distribution of errors, analogous to the transposition gradient in the serial-order literature. More specifically, I examined whether the neighbor effect observed in the first study is a topological phenomenon, reflecting an effect of neighborhood, or whether it reflects an effect of metric spatial proximity in WM, such that closer items are confused more often than those further apart (spatial transposition gradient). To answer this question, I used visual recognition and reconstruction tasks and manipulated the distance between memory stimuli and examined whether the likelihood of transposition errors depends on the spatial distance between memory stimuli. The results showed that transposition errors occurred significantly more often than predicted by chance and that extra-list errors (i.e., reporting an item that has not been part of the memory set) occurred less often than predicted by chance. Transposition errors followed a spatial gradient, that is, the smaller the metric distance between representations, the more likely they were confused with each other, in line with the prediction of a spatial transposition gradient. The emergence of this gradient can be explained by overlapping retrieval cues, analogous to the explanation for the transposition gradient in serial recall. Accordingly, context cues overlap with each other as a function of spatial proximity and as a consequence, bindings are not only established to the corresponding content, but as well to other items (although weaker) as a function of spatial distance between representations. When attempting to retrieve an item by its spatial position, these misleading bindings cause the partial retrieval of other WM contents and thereby frequently result in transposition errors.

In sum, the study provided evidence for a spatial transposition gradient, if retrieval cues are spatial position. This finding shows the relevance of cue-based retrieval in WM and

contributes to the unity of WM mechanisms by showing the limited distinctiveness of WM representations in the spatial analogous to similar findings for the temporal dimension.

2 Discussion

The first study showed that non-focused information remained accessible in WM even if attention is focused on another representation and even if this non-focused information is known to be irrelevant for the remainder of the task. De-focused information remained highly accessible and identifiable after retracting focal attention from it. The second study revealed a spatially specific distribution of transposition errors in WM, that is, the closer representations are in WM, the more likely they are confused with each other.

2.1 The Explanatory Power of the Three WM Models

Different predictions can be derived from the three WM models regarding the investigated phenomena. These diverging expectations can be utilized to evaluate the appropriateness of the different WM conceptualizations underlying these models. In the following, I will discuss how well each of the models can account for the data and in what regards the models could be modified to give a better account.

2.1.1 The Fate of Non-focused and De-focused Information in WM

The finding that focusing attention on a representation does not impair the maintenance of non-focused information is in line with the assumption of separate mechanisms for focusing and for maintaining information in WM, which do not draw from the same limited capacity. This assumption is incorporated in the design of the three-embedded-components model (Oberauer, 2002, 2009). In contrast, the slot and the resource model both embody the assumption of a unitary capacity limit for the maintenance and the focusing of WM representations and therefore predict that the maintenance of non-focused information will be compromised by focusing attention elsewhere, a prediction that is in conflict with the results obtained.

None of the models makes explicit predictions for the fate of de-focused information in WM and none of them seems to be capable of accounting for the elevated accessibility of de-focused information in their current versions.

Within the slot model, this high accessibility of de-focused information could be explained by assuming that the extra-slot (stemming from the limited number of slots) allocated to the representation during focusing remains on that representation even after de-focusing it. However, the model assumes that focusing and maintaining WM representations draw from the same limited number of slots, predicting trade-offs between these two processes. Hence, if extra-slots are allocated to the focused and the de-focused representation, the number of remaining slots for maintaining non-focused information would be severely reduced. In light of the typically estimated capacity limit of three to four representations (e.g., Cowan, 2005; Zhang & Luck, 2008), allocating two slots to the focused and remaining another two slots on the de-focused information implies that no slot is left to maintain non-focused information in WM. As a consequence, if participants have to de-focus again and focus a third representation, performance for this third representation should be dramatically impaired and drop to chance level. A similar prediction can be derived from the resource model, which - as the slot model - relies on the assumption that focusing and maintaining representations rely on the same mechanisms. However, according to the resource model, the performance drop is expected to be less dramatic than expected on the basis of the slot model, because activation can be allocated continuously, which could prevent a scenario in which no activation is left to maintain non-focused information. These predictions are in conflict with the results, which showed that three representations can be focused sequentially without catastrophic forgetting, leading to similar performance as focusing two representations sequentially.

The resource model further incorporates the assumption that separate resources are responsible for the storage of *contents* (e.g., color) and the storage of *content-context bindings*

(e.g., color at a spatial position) (Bays et al., 2011). Whether contents or bindings are strengthened by focusing has different implications for the fate of de-focused information. Although a recent study (Bays et al., 2011) indicated that the focusing benefit relies on the allocation of extra-activation to the *content* (improved precision due to focusing) and does not include extra-activation of the binding (transposition errors were unaffected by focusing), our findings cannot be accounted for by assuming that only the content is strengthened by focusing: If the extra-activation from focusing remains on the de-focused *content* and attention is re-focused on another representation, both representations are highly activated at the same time, but not more identifiable, because the bindings, which are needed to recollect the representation's context, are not strengthened in this scenario. Hence, these highly activated representations should strongly compete for selection and as a consequence, recognition of de-focused information should be impaired, contrary to our finding. Otherwise, if the extra-activation remains on the de-focused *content-context binding*, performance on de-focused representations should be improved compared to non-focused information due to being highly accessible and identifiable by its context information (spatial position), in line with the results obtained in our study.

As the slot and the resource model, the three-embedded-components model does not make specific predictions for the fate of de-focused information. Nevertheless, it could be expected that de-focused information is indistinguishable from other information in the region of direct access, because the focus of attention is functionally limited to a single representation (the currently focused one). Contrary to this expectation, the results showed that de-focused information remained highly accessible and discriminable. Although the model does not predict this, it can account for this finding by an additional assumption, namely that focusing a representation strengthens the binding between content and context not only for the time during which it is focused, but even after attention is shifted elsewhere. By

this assumption, de-focused information can be maintained highly accessible in the region of direct access.

In sum, the three-embedded-components model, which assumes that focusing and maintaining WM representations rely on separate mechanisms, can account for the maintenance of non-focused information when attention is focused elsewhere. The other two models, which assume that focusing and maintenance draw from the same capacity limit, predict trade-offs between focusing and maintaining information; a prediction that is not supported by the data. None of the models can readily account for the data pattern of de-focused information. However, the problems the models face differ: Both the slot and the resource model exhibit basic conceptual issues, namely the assumption that focusing and maintaining WM representations relies on the same mechanisms. The three-embedded-components model, which relies strongly on bindings and incorporates separate mechanisms for focusing and maintaining, can be extended by the assumption that de-focused information's content-context bindings remain strengthened even after focused attention has been retracted from it to account for the high accessibility of de-focused information.

2.1.2 The Spatial Distribution of Transposition Errors

None of the three considered models predicts the spatial transposition gradient and none of them can account for this finding without additional assumptions.

Although the slot model provides a mechanism to account for the distribution of errors according to a precision gradient (i.e., the more similar two items are, the more likely they are confused), it incorporates the assumption that items are either stored in a slot or are completely lost. As a consequence, errors are assumed to be distributed randomly (Zhang & Luck, 2008, 2011). This is in conflict with the spatial transposition gradient the data revealed. However, the model could be extended to account for the gradient by assuming that the spatial location serves to select the corresponding slot: When getting the retrieval cue, which

indicates which location will be probed, the corresponding slot has to be identified by finding the slot that matches this information best. In this stage, transposition errors depending on a spatial gradient can occur. In the next stage, the content (e.g., color) of the selected slot has to be retrieved and in this step, errors can occur as a function of color similarity, as predicted within the slot model (Zhang & Luck, 2011).

The high proportion of transposition errors relative to guessing errors can be accounted by the resource model by the assumption that noise corrupts bindings in WM (Bays et al., 2009), but no specific distribution among transposition errors is predicted. The spatial transposition gradient can be explained within this model by assuming that misleading bindings are established, which associate a retrieval cue (spatial position) to other contents than the corresponding one as a function of spatial distance between representations: The closer items are to each other, the stronger misleading bindings are available and the more likely are transposition errors at retrieval.

The three-embedded-components model does not specify predictions for the distribution of transposition errors. However, because bindings are already central in the context of this model (cf. Oberauer, 2005a), it can be extended by one additional assumption - similar to the above proposed extension for the resource model - to account for the gradient: The strength of misleading bindings depends on the spatial distance between items with stronger misleading bindings for closer representations. By this adaptation the model incorporates the spatial transposition gradient without compromising its core conceptualization of WM.

In sum, none of the models can account for the spatial transposition gradient in its current version. The slot model explicitly predicts that errors in WM should be distributed randomly, while the resource model can account for the high prevalence of transposition errors, but does not incorporate the assumption of a specific distribution among these errors. As the resource model, the three-embedded-components model does not specify predictions

for the distribution of transposition errors, but could account for the gradient by the additional assumption that the strength of misleading bindings is determined by a function of the spatial distance between WM representations, reflecting limited spatial distinctiveness of WM representations.

2.2 Two Important Characteristics of WM

The two studies revealed two important characteristics of WM, namely the relevance of bindings, and the assumption of separate mechanisms for the maintenance and for focusing representations in WM.

2.2.1 Bindings in WM

Bindings are a powerful mechanism to account for a variety of WM effects, including also the effects of focusing and de-focusing WM representations and the spatial distribution of transposition errors (by misleading bindings). In the two studies, two factors were identified, which significantly affect the strength of these bindings: Attentional focusing and spatial proximity of WM representations. Focusing attention on a single representation in WM increases the strength of the binding between the content and its context. This renders the representation more accessible and discriminable from other WM information and thereby facilitates retrieval and improves performance, reflected in a focusing benefit. The bindings remain strengthened even after focusing another representation, leaving the de-focused representation highly accessible. As a consequence, when encountering a de-focused representation, its identity can be retrieved by the bindings and behavior can be guided by that knowledge. Additional support for the importance of bindings in WM stems from the finding of a spatial transposition gradient, when the spatial position of an item serves as a retrieval cue. The spatial overlap between representations' context cues influences the extent to which misleading content-context bindings are established: Misleading bindings, which associate a context cue to other contents than the corresponding one, are established as a function of the

spatial distance between representations and this can account for the spatial transposition gradient.

2.2.2 Separate Mechanisms for Focusing and Maintaining WM Representations

The assumption that focusing and maintaining information in WM are not accomplished by the same mechanisms, not drawing from the same limited supply (e.g., a limited number of slots or a limited resource such as activation) is crucial to account for the finding that focusing attention on a representation does not compromise the maintenance of non-focused and de-focused information. The assumption of separate mechanisms for the two processes is incorporated in the three-embedded-components model, which can account for the fate of non-focused information while attention is focused elsewhere. Furthermore, this assumption is important to account for the finding that de-focused information remains highly accessible, while another representation is focused and the maintenance of non-focused information is not comprised. In contrast, models that assume that maintenance and focusing rely on the same mechanisms have difficulties to account for this pattern.

2.3 Future Directions

In this thesis, I have addressed questions that have received little attention in WM research and in theorizing so far. The results obtained in the two studies provide new insights into the organization of WM and have implications for several prevailing WM conceptualizations, which cannot (fully) account for the data. However, several other related questions emerged in the course of this work. In the following, some of these open questions are introduced and possible answers to them are sketched.

2.3.1 Spatial Imprecision of Retrieval Cues

The second study showed that item-context bindings appear to be spatially imprecise, as an increasing metric distance between two representations decreases the chance of

confusing them. A question that is still open is whether there is a difference between directly neighboring items and non-neighboring items, when their spatial distance to the target is the same. If the distribution of transposition errors relies solely on the metric distance between items, there should be no difference between the two scenarios, because the metric distance is identical in both cases. Alternatively, if it is not purely the metric distance that affects the distribution of errors, having an intervening item between the target and the non-neighbor might alter performance, as misleading bindings to the non-neighbor might be weakened by stronger (misleading) bindings to the direct neighbor, which sits between the non-neighbor and the target item. In this sense, intervening items could attenuate the spatial gradient for non-neighboring items. In future research the effects of intervening items on misleading content-context bindings should be examined in more detail.

2.3.2 A Limit of Strengthened Bindings?

Focusing appears to strengthen the corresponding representation's content-context binding, which remains strengthened even after attention has been retracted and focused on another representation. However, it is unclear, how many bindings can be strengthened and maintained strengthened at a time and how these bindings are disengaged. As proposed in the three-embedded-components model, the region of direct access is limited regarding the amount of bindings that can be maintained simultaneously. This limit is estimated to be about three to four, hence, within the range of strengthened bindings tested in the first study. Several scenarios can be hypothesized if more bindings need to be focused and strengthened: First, there might be a limit up to which bindings can be strengthened, and after exceeding this limit, no more bindings are strengthened. Second, older strengthened bindings might be weakened or dissolved to make room for new strengthened bindings. Third, by strengthening bindings and thereby improving these representations' discriminability, interference between these bindings might be reduced and thereby the limit regarding the amount of bindings,

which can be maintained simultaneously in the region of direct access, might be boosted. Forth, the ability to maintain strengthened bindings might be limited by time, such that binding strength decreases as time passes. Future research will have to unravel which processes are engaged in the decomposition or dissolution of strengthened bindings.

2.3.3 Computational Modeling

The examination of the verbal versions of the slot, the resource and the three-embedded-components model revealed that none of them can account for all of the findings, and I invoked some processes and mechanisms, which could be responsible for the pattern of effects observed in the data. However, although these processes seem to be adequate to produce the effects observed, only computational modeling can determine how these processes affect each other and other processes, which are incorporated in the models already, and whether they can predict the pattern observed in the data. Computational modeling provides a powerful tool to quantitatively evaluate the explanatory and predictive power of models to account for data. All models examined here have been implemented as computational models (Bays & Husain, 2008; Oberauer, Souza, Druey, & Gade, 2012; Zhang & Luck, 2008, 2009). A next step could be to implement the processes proposed to account for the data into these existing versions of the models and to evaluate, whether these adaptations lead to predictions in line with the empirical findings without altering other, valid predictions of the models. This can provide further insights into the processes underlying our findings and broaden the understanding of WM in general.

2.4 Conclusion

This thesis revealed two important characteristics of WM: First, the explanatory power of bindings to account for fairly diverse phenomena in WM, such as the fate of focused, non-focused and de-focused information in WM and the spatial distribution of transposition errors; and second, the importance of separate mechanisms for focusing and maintaining

representation in WM. These results allowed the evaluation of competing WM models, namely, the slot, the resource and the three-embedded-components model, and the identification of weaknesses in their conceptualization of WM, indicating that all these models need to be modified – though to different extents – to account for the findings.

Part II: Empirical Studies

3 Focused, Unfocused, and De-focused Information in Working Memory

Laura Hein and Klaus Oberauer

University of Zurich

Submission status:

Resubmitted after revision

Authors' contributions:

Laura Hein: Literature review, development of the research question and goals, design of the study, programming experiments in MATLAB, data collection, analysis, and interpretation, writing and submission of the manuscript

Klaus Oberauer: Supervision and discussion Laura Hein's contributions and revision of the manuscript

Acknowledgements:

This research was funded by the Swiss National Science Foundation (project number 100014_126766 / 1). We thank Matthias Hartmann, Eszter Montvai, and Iliana Karipidis for their assistance in the data collection.

Abstract

The study investigates the effect of selection cues in working memory (WM) on the fate of not-selected contents of WM. Experiments 1A and 1B show that focusing on one cued item in WM does not impair memory for the remaining items. The non-focused items are maintained in WM even when this is not required by the task. Experiments 2 and 3 show that items that were once focused in WM remain strengthened after the focus shifts away from them. When de-focused items are presented as mismatching recognition probes, they are rejected better than other mismatching probes (Experiments 2 and 3). When a de-focused item was later cued again, such that the focus had to shift back to it, then that item was recognized better than an item cued for the first time (Experiment 3). The results support the distinction between mechanisms for temporary maintenance and the focus of attention in WM, and challenge theories that explain maintenance and focusing by the same mechanisms, such as a limited number of slots or a limited resource.

Keywords: working memory, attention

3.1 Introduction

Often we have to hold several pieces of information in working memory (WM), but for a particular cognitive operation we need to focus only on a subset of them. What are the consequences of focusing for the remaining, unfocused information in WM, and what happens with the focused information once it is de-focused later? In this article we investigate how focusing information affects non-focused and de-focused information in WM for visual information. We interpret the results within three theoretical frameworks for characterizing WM: Theories assuming a discrete WM capacity (Cowan, 2005; Zhang & Luck, 2008), constant-resource theories (Just & Carpenter, 1992; Bays & Husain, 2008), and the three-embedded components framework by Oberauer (2002, 2009).

3.1.1 Discrete-Capacity Theories

Cowan (2001) introduced the hypothesis that WM has a discrete capacity of approximately four items or chunks. This capacity limit applies to the central component of Cowan's embedded-processes model of working memory (Cowan, 1988, 1995, 1999). The model assumes two embedded structural components contributing to WM: the activated part of long-term memory (peripheral component) and the focus of attention (central component). The activated part of long-term memory facilitates retrieval of potentially relevant representations in long-term memory; it has no capacity limit but is limited by decay and interference. In contrast, the focus of attention is assumed to have a limited capacity, such that there is a maximum number of independent representational units (i.e., chunks) that can be maintained in the focus at any one time.¹ This capacity limit is conceptualized as a discrete,

¹ The capacity limit according to Cowan (2001) applies to chunks, which are learned units that can consist of several list items in an experiment. In the present context, each item of the memory set can be regarded as a separate chunk, and therefore we will from here on just speak of items or (in the case of visual WM) of objects.

fixed-capacity limit of about three to four units in healthy young adults (Cowan, 2001, 2005; Cowan, Rouder, Blume, & Saults, in press).

Zhang and Luck (2008, 2009) formalized the notion of a discrete capacity as a simple mathematical model for visual WM, called the *slot model*. The slot model proposes that chunks in WM are stored in an all-or-nothing fashion up to an individual limit, which is set by the number of available slots in WM. Only a single chunk (e.g., one object with all its features) can be stored in each slot. If the number of items to be stored exceeds the capacity limit, no more slots are available for storage and further items are lost to WM; consequently, participants can only guess about their identity, and retrieval of those items decreases to chance level.

As a refinement of the slot model, Zhang and Luck (2008) proposed the *slot-and-averaging model*. According to this model, multiple copies of an item can be stored in multiple slots if more slots are available than there are items competing for them. Storing an item in more than one slot improves the precision with which that item's features can be retrieved, because information about this item from multiple slots is averaged, thereby reducing the variance of stored feature information. Assigning an item to multiple slots can also be used to prioritize that item; this provides a mechanism for focusing on an item within a set of items in the central component of WM. Because focusing within WM is the focus of our present work, we will use the slot-and-averaging model to derive predictions from discrete-capacity theories for our experiments.

3.1.2 Constant-Resource Theories

Another conceptualization of the capacity limit in WM is to assume that a limited resource is responsible for the observed limit. Constant-resource theories have a long tradition in research on verbal WM (e.g., Just & Carpenter, 1992); more recently, a constant-resource theory has been proposed as an alternative to the slot model for visual WM (Bays & Husain,

2008). The idea is that a constant limited resource (e.g. of activation) can be flexibly distributed across the contents of WM.²

By default, every representation can be assumed to receive an equal share of this resource (Bays & Husain, 2008). With increasing number of items to be stored, each of them receives a smaller share of the resource, resulting in decreased accuracy of retrieval.

However, representations in WM are not necessarily all activated to the same extent. A subset of them can be focused by receiving more activation than others. Importantly, the overall activation remains constant. Therefore, if some representations are prioritized by receiving extra activation, the activation of the remaining items must be decreased accordingly.

3.1.3 The Three-Embedded-Components Theory of WM

The three-embedded components framework by Oberauer (2002, 2009) builds on the embedded-processes model (Cowan, 1995, 2001). The three embedded components can be understood as successive levels of selection of memory representations for (cognitive) action. The broadest level of selection is the activated part of long-term memory. It consists of the information activated that is potentially relevant for the task and likely to be required later on, similar to Cowan's (1999) conceptualization of the activated part of long-term memory. The second component, the region of direct access, consists of a subset of representations in the activated part of long-term memory which is temporarily bound to context cues through which they can be accessed. The region of direct access is similar to Cowan's focus of attention in that it is the central, capacity-limited component of WM, which can hold up to about four chunks simultaneously. Different from Cowan's theory, however, the capacity

² Cowan et al. (2012) note that the slot-and-averaging model introduced by Zhang and Luck (2008) takes an intermediate position between strict discrete-capacity theories (such as the slot model) and resource theories. They propose further possible intermediate assumptions, such that a limited resource that can be freely allocated up to a maximum number of elements. For the present experiments, these intermediate conceptualizations make the same predictions as the constant-resource hypothesis, and we therefore do not consider them separately.

limit of the direct-access region is not a fixed number of chunks. Rather, the capacity limit arises from interference between item-context bindings, and therefore the number of chunks that can be accommodated in the direct-access region depends on the discriminability of the items and their contexts (for details see Oberauer, Lewandowsky, Farrell, Jarrold, & Greaves, 2012).

Within the region of direct access, individual items must often be selected as the target of the next cognitive operation. These items are held in the focus of attention, the narrowest level of selection in WM in the three-embedded components framework. This focus is assumed to temporarily select a single item or chunk in the region of direct access. Note that the focus of attention in the theory of Oberauer (2002, 2009) differs from the focus of attention in the theory of Cowan (1999, 2005), in that it selects a single item within the set of items currently held in the central component of WM.

The activated part of long-term memory has a similar role in Cowan's and in Oberauer's model: It consists of activated representations which might be relevant for the task at hand. Activated long-term memory is unlikely to make a substantial contribution to the task we used in this study, in which we required the recognition of briefly presented colored circles. Performance in this procedure is limited to approximately three items (Makovski, 2012; Makovski, Sussman, & Jiang, 2008; Kuo, Stokes, & Nobre, 2012), which is, if anything, less than expected from the "magical number four" that Cowan (2001, 2010) proposed as an estimate of the central component of WM. Moreover, there appears to be little proactive interference in this type of task (Lin & Luck, 2012).

To prevent confusions due to the similar terminology in the models by Cowan (1995) and Oberauer (2002, 2009), we will refer to Cowan's focus of attention and Oberauer's region of direct access by the theory-neural term *central component of WM*, throughout the remainder of this article. The term *focus of attention* will be used for designating the narrow focus of attention in the three-embedded-components theory (Oberauer, 2002).

3.1.4 Focusing Items within Central Working Memory

There is converging evidence for the notion that people can prioritize individual items within central WM. Experiments involving sequential encoding or processing of information in WM have shown that the one item last encoded or last used is accessed particularly quickly and accurately (Garavan, 1998; McElree, 2001; Oberauer, 2006, Oberauer & Bialkova, 2009; Woodman & Vecera, 2011). For instance, repeatedly updating the same item leads to faster reaction times than updating two different items in succession (Garavan, 1998; Oberauer, 2002). Likewise, repeatedly retrieving the same objects from WM to report different features of that object is easier than switching between objects (Woodman & Vecera, 2011).

Several studies have investigated the beneficial effects of focusing attention in WM in response to a cue (e.g. Griffin & Nobre, 2003; Landman, Spekreijse, & Lamme, 2003; Makovski & Jiang, 2007, 2008; Makovski, Sussman, & Jiang, 2008; Nobre, Griffin, & Rao, 2008; Sligte, Scholte, & Lamme, 2008). Typically, variations of a visual short-term recognition paradigm (Luck & Vogel, 1997) are used in those experiments. In these tasks participants are required to memorize visual stimuli presented simultaneously in a memory display and compare them (or parts of them) to a probe display. So-called *retro-cues* are presented in the interval between the offset of the memory display and the onset of the probe display, while the items are held in WM. These cues indicate which part of information from the memory display has to be compared to the probe display, thereby directing attention to specific WM contents. Using valid retro-cues leads to performance improvement compared to providing the same task with uninformative cues (Griffin & Nobre, 2003; Lepsien, Griffin, Devlin, & Nobre, 2005; Nobre et al., 2008) or with invalid cues (Griffin & Nobre, 2003). This effect is referred to as the *retro-cue benefit*. Because the cue is presented after the memory display, the focusing benefit cannot be explained by a neglect of non-cued items during encoding. Rather, it seems that attention can be applied to a single WM representation and that focusing on specific information in WM increases its accessibility.

In the experiments reported here we use the retro-cue paradigm to induce people to focus a single object within a set of several visual objects held in central WM. We next consider how the three theoretical frameworks introduced above can be applied to the focusing benefit.

3.1.5 Mechanisms of the Focusing Benefit

All three theories are compatible with the retro-cue benefit, but they differ in their predictions of the consequences of focusing on the non-focused (non-cued) representations within the central component of WM. According to the discrete-capacity theory, the retro-cue benefit can be explained by the assumption that the focused item is assigned to more than one slot, thereby gaining in precision through averaging (Zhang & Luck, 2008, 2009). That is, instead of assigning each object to a single slot, at least two slots are given to the focused object. As a consequence, however, fewer slots are available to the remaining, non-focused information. Therefore, for memory displays exceeding the capacity, the fixed-capacity theory predicts that performance for non-cued information suffers from focusing another representation. In other words, the retro-cue benefit implies a retro-cue cost for non-focused information.

The constant-resource framework assumes that activation of WM contents can be flexibly allocated to single objects held in the central component, but that activation is a limited resource within this component. According to this theory, the cueing benefit can be explained by assuming that more activation is allocated to the cued item than to the other items (Bays, Gorgoraptis, Wee, Marshall, & Husain, 2011). Thereby, its likelihood of being correctly retrieved is increased compared to the other items. Crucially, when the activation of the focused item is increased, less activation is available for the remaining items in central WM. Hence, the retro-cue benefit again implies a cost for non-focused information in the

central component of WM, resulting in the same prediction for non-focused information as the fixed-capacity theory.

In the three-embedded-components theory (Oberauer, 2002, 2009), the focus of attention is separate from the capacity-limited central component of WM (i.e., the direct-access region). According to this theory, the retro-cue benefit arises because the cued object becomes the content of the focus of attention. Selecting an object into the focus enhances its accessibility, but does not impair the representations of other objects in the direct-access region. Therefore, in contrast to the other two theories, no detrimental effect of the retro-cueing benefit is predicted for the non-focused representations.

In sum, focusing information in WM prioritizes the focused item and causes a performance benefit if this information is required for the next cognitive action. The models do not differ regarding their predictions for the focused item, but propose different fates for the remaining WM contents. The present study asks what happens to them. We distinguish non-focused information -- information that has not been prioritized during the retention interval -- and de-focused information -- information that had been recently prioritized, but then replaced by other information to be prioritized. Examining the fates of these kinds of information will provide information for evaluating the models introduced above. In the following section we detail the hypotheses we test in the present experiments, and summarize the scant available evidence on the effects of focusing on the fate of non-focused and de-focused information.

3.1.6 Non-Focused and De-Focused Information in WM

Non-Focused WM Contents

Non-focused information refers to information inside central WM that has not been focused (attended) before after initial encoding. As discussed above, fixed-capacity theories and constant-resource theories of WM lead to the prediction that, as one item is focused, other

items in central WM tend to be forgotten more, compared to a situation where no item is focused. In contrast, the three-embedded-components theory predicts that focusing one item has no effect on memory accuracy for the remaining items.

Support for increased forgetting of non-focused information in WM was provided by a series of experiments by Matsukura, Luck, and Vecera (2007). The authors compared two different accounts for the beneficial effect of focusing in WM: the protection and the prioritization account. The protection account assumes that focused attention protects the cued item from degradation and interference, while the non-focused items decay or suffer from interference. In contrast, the prioritization account presumes that the cued item is given priority in the comparison process, while non-cued items remain available for later comparison. The authors presented retro-cues pointing either to the left or the right hemisphere of the display in a visual WM paradigm. In a subset of trials, two successive cues were presented; the last of them was always valid. Two conditions were distinguished for the two-cue trials: Either both cues were pointing to the same hemisphere (e.g. set of items on one side of the fixation cross) or the two cues pointed in opposing directions. When both cues pointed in the same direction, memory performance was as good as in single-cue trials. When the second cue pointed into another direction than the first cue, performance was worse than in single-cue trials. The authors interpreted this as support for the protection account, assuming that while participants focus on the hemisphere cued by the first cue, the information in the other hemisphere became prone to forgetting. This forgetting could be due to resources or slots being taken away from the non-focused items, but it could also be due to decay or interference acting on the non-focused items during the time of the second cue. According to the latter interpretation, which was endorsed by Matsukura et al. (2007), it is not the act of focusing itself that impairs memory for the non-focused information. Therefore, their result does not adjudicate between the three theories considered here.

Evidence for unimpaired maintenance of non-focused information was provided by a study by Landman et al. (2003) using a change detection task. Change detection is a variant of the short-term recognition paradigm in which participants compare a memory display to a probe display and determine whether they match, or whether there was a change from the first to the second display. Landman and colleagues presented either one or two successive cues in the retention interval, indicating which single item from the memory display might change from memory display to probe. In the case of two cues, the last cue (which pointed in a direction other than the first cue) was always valid in predicting the location of the possible change. Additionally, onset time of the valid cue was manipulated (early vs. late). Results revealed no difference between early and late valid cue onset, and memory performance was not impaired by the first cue in two-cue trials compared to single-cue trials. These results indicate undiminished maintenance of non-focused objects and flexible re-allocation of attention, supporting the assumption that non-focused information is retained in memory. Landman et al. (2003) presented the memory display and the two cues in rapid succession, so that it is not clear to what extent sensory memory contributes to performance.

So far, we have only considered two possible fates of non-focused information in WM, maintenance and forgetting. A third possibility is that non-focused items are actively removed from central WM because they are no longer needed. An assumption of the three-embedded-components model is that information no longer needed is removed from central WM, thereby reducing the interference that limits the capacity of central WM (Oberauer, 2002, 2005b). Evidence for removal of information comes from experiments using a modified Sternberg task. In these experiments, two memory lists were presented for encoding, and a retro-cue indicated which of the two lists was relevant for a recognition test. Reaction times (RTs) for responses to the recognition probe revealed robust setsize effects for the relevant list, whereas the setsize effect of the irrelevant list disappeared with increasing time between the cue and the test stimulus. One to two seconds after the cue, the irrelevant setsize had no

influence on RTs. This time reflects the time it takes to remove non-focused information from the central component of WM. Removed information, however, was not lost, but could be retrieved from activated long-term memory: When a second cue pointed to the previously irrelevant list, the setsize effect for that list re-appeared in RT data for the second recognition probe (Oberauer, 2005b; cf. Lewis-Peacock, Drysdale, Oberauer, & Postle, 2012). These data indicate that WM contents are under flexible control and can be removed from the central components of WM if they become temporarily irrelevant.

Active removal can be distinguished from passive forgetting because the former, but not the latter, is under flexible control, depending on task demands. Passive forgetting would affect non-focused information regardless of its relevance, whereas a person would actively remove information only if it is perceived as no longer relevant. In Experiments 1A and 1B, we test whether people can flexibly control the fate of all non-cued items by manipulating whether non-focused information is rendered irrelevant or remains relevant.

De-Focused WM Contents

The status of non-focused information might differ from the status of previously focused and then de-focused information. Both types of information are in central WM and currently not focused, but differ regarding their focusing history. Non-focused information has never been focused after encoding. In contrast, de-focused information has recently been focused, but is currently not focused anymore. Little research has been done on the status of de-focused WM contents, and their fate has not been considered in the three WM theories introduced above. We therefore cannot derive predictions for de-focused information from these theories. Nevertheless, three broad hypotheses regarding its fate can be contrasted: de-focused information can be set back to baseline level of memory strength, it can be inhibited, or it can be strengthened compared to non-focused WM representations. In the context of

Experiment 2 we will elaborate these hypotheses, and in the General Discussion we will discuss how our results can be accommodated by the three theories.

The present study investigates the fate of non-focused and de-focused information in WM for visual information. Experiments 1A and 1B provide support for the maintenance of non-focused information in WM, even if this information is not required for the task. Experiments 2 and 3 investigate the fate of de-focused information and show that de-focusing leaves the respective item strengthened relative to information that was not focused during the retention interval.

3.2 Experiment 1A and 1B: Retro-Cue Benefit and Non-Focused Information in WM

The main aim of both Experiments 1A and 1B was to examine how prioritizing one item in WM affects memory for the remaining items. Furthermore, in Experiment 1A, we intended to replicate the retro-cue benefit in our task, that is, the finding that cues orienting attention to one item in WM improve retrieval accuracy for that item. For these purposes, we used a short-term color recognition task, which required participants to memorize a multi-item memory display and compare one of these items to a single-item probe display. The experimental design was modeled after the third experiment described in Landman et al. (2003): After encoding the memory display, either one or two successive cues were presented. Participants were correctly informed that the last cue was always valid. Our experiments differed from Landman et al.'s (2003) experiment in three regards: First, we increased several time intervals to examine focusing effects on non-focused information in WM – over a time course that unambiguously excludes contributions from iconic memory. Second, in Experiment 1A, we added two no-cue conditions, serving as baselines to evaluate whether the retro-cues were beneficial to memory performance. Third, we examined whether the fate of non-focused WM representations is under flexible control by manipulating the predictability

of the number of retro-cues as a within-subject variable in Experiment 1A, and as a between-subject variable in Experiment 1B.

In line with the well-established retro-cueing benefit (e.g. Griffin & Nobre, 2003; Lepsien et al., 2005; Nobre et al., 2008), we expected cue trials to show better performance than no-cue trials in Experiment 1A. To test whether non-cued information remains accessible after attention was cued to another location, we compared performance in single-cue trials and two-cue trials. If non-cued items are maintained in central WM, and people can flexibly shift the focus of attention within it, we predict that performance on two-cue trials should be as good as in single-cue trials. In contrast, if focusing one item is detrimental to memory for the remaining non-focused items, performance in single-cue trials should be better than performance in two-cue trials.

To assess whether people have control over whether they maintain or remove non-cued information, we created two predictability conditions. In the predictable condition there were no two-cue trials, so that after seeing the first cue, participants knew that this cue validly points to the item that will be tested. Therefore, participants could remove non-focused information from the central component of WM after receiving a cue, thereby reducing memory load and improving task performance. In the unpredictable condition, single-cue and two-cue trials were mixed randomly, so that after seeing the first cue, participants did not know whether this cue would be valid in the end. In this condition, it would be unwise to remove all non-cued items upon seeing the first cue, because doing so would jeopardize performance in two-cue trials. Therefore, in the unpredictable condition participants should rather maintain non-focused information after the first cue. If participants followed their optimal strategy under each predictability condition, this would result in higher WM load in the unpredictable condition than in the predictable condition. Hence, comparing performance on single-cue trials between predictable and unpredictable conditions reveals whether non-focused information can be flexibly maintained or removed according to task requirements.

To summarize, we can distinguish three possible outcomes: If non-focused information is *maintained* in WM, we should observe no difference between single- and two-cue trials, as well as non-distinguishable performance between predictability conditions. In case non-focused information is *weakened* or entirely lost, performance in single-cue trials should be better than in two-cue trials regardless of predictability. Finally, in case of *flexible control*, performance in the predictable condition should be better than in the unpredictable condition on single-cue trials, and single-cue and two-cue trials should yield equal performance in the unpredictable condition. A summary of predictions can be found in Table 1.

Table 1

Predictions for non-focused information in WM (Experiment 1A and 1B)

	Predictions for non-focused information		
	Maintenance	Weakening	Removal (flexible control)
Effect of Number of Cues	2-cue = 1-cue	2-cue < 1-cue	2-cue = 1-cue
Effect of Predictability	Pred = Unpred	Pred = Unpred	Pred > Unpred

Note. Overview of the predictions from three hypotheses (columns) for the effect of number of cues and predictability of the number of cues in Experiments 1A and 1B. Pred = predictable condition; Unpred = unpredictable condition. If non-focused items are always maintained in working memory, no difference between single-cue and two-cue trials and between predictable and unpredictable conditions is predicted. If non-focused items are weakened by focusing one item, two-cue trials should be worse than single-cue trials. Again, no effect of predictability is expected. If maintenance of non-focused items is under flexible control, there should be no difference between single-cue and two-cue trials (in the unpredictable condition), and performance (on single-cue trials) should be better in the predictable than the unpredictable condition.

Many studies investigating WM with visual materials asked participants to engage in articulatory suppression to ensure that they relied exclusively on visual, as opposed to verbal, WM. In this and the following experiments we did not use articulatory suppression for two reasons. First, we are interested in general mechanisms of WM, rather than mechanisms that apply only to WM for purely visual representations. Therefore, it is not essential to our conclusions that people relied exclusively on visual representations. Second, several studies

have shown that controlling for verbal re-coding of visual material in tasks similar to ours does not affect performance (e.g. Luck & Vogel, 1997; Morey & Cowan, 2004). Therefore, there was little to be gained from placing the extra burden of articulatory suppression on participants.

3.2.1 Method

Participants

Twenty-four students from a Swiss University participated in Experiment 1A. Their mean age was 24 years (range: 18-30) and 4 of them were male. Participants received financial incentives (30 CHF for approximately two hours) or partial course credit in exchange for their participation. For Experiment 1B, sixty participants (30 per group, 80% female) were tested with a mean age of 23 years (range predictable group: 18-31; range unpredictable group: 18-33).

Task and stimuli

All participants were tested individually in a laboratory cabin. All experiments were programmed with the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) implemented in MATLAB.

Experiment 1A. At the beginning of each trial, a memory display was shown for memory encoding. The memory display consisted of six colored circles (randomly chosen from a set of nine colors), arranged on an imaginary circle centered on a black screen. The display remained on the screen for 1 s and was followed by an interval (see Figure 1 for interval lengths) during which the screen went black. After this interval, none, one, or two retro-cues were presented sequentially with an inter-cue interval of 500 ms.

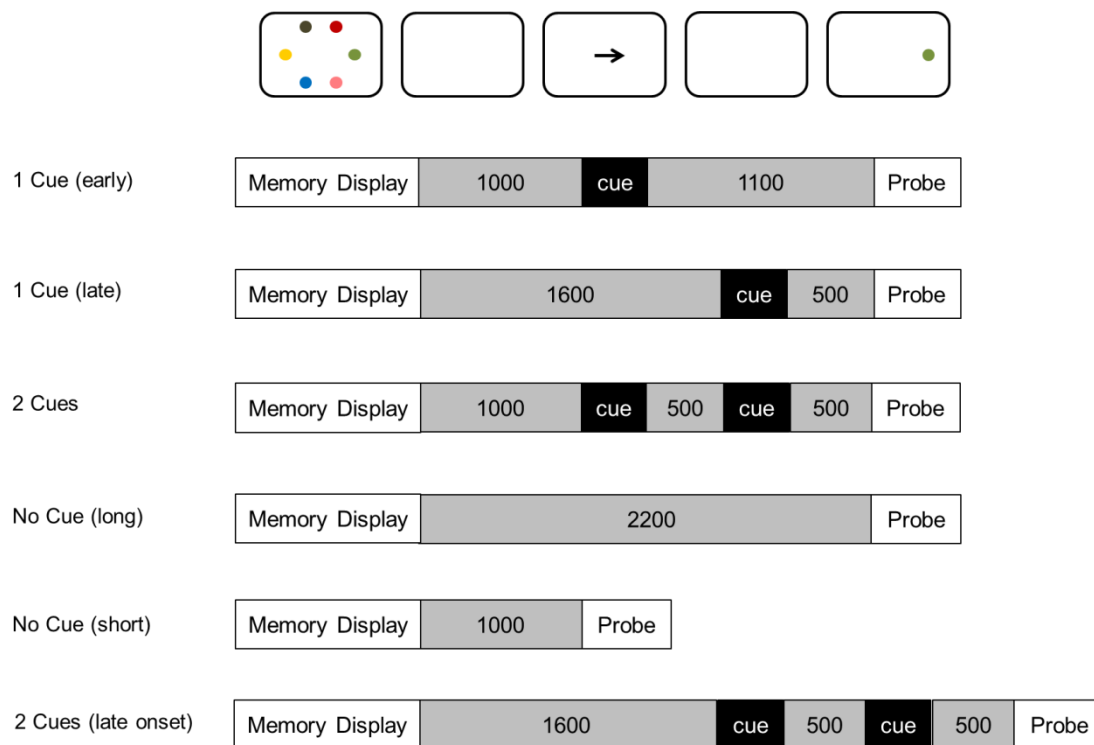


Figure 1. Overview of the four experimental conditions in Experiment 1A (upper five rows) and of Experiment 1B (upper three rows and last row). The memory display was presented for 1 second, and the probe display remained on screen until the participant entered the answer via the keyboard. Cues were presented for 100 ms. The interval lengths in ms are depicted in the corresponding intervals. Two predictability conditions were distinguished. In Experiment 1A the predictable condition comprised both single-cue and no-cue trial condition (rows 1, 2, 4, and 5) and the unpredictable condition comprised all conditions of rows 1 to 5). In Experiment 1B, the predictable group comprised single-cue trials (upper two rows) and the unpredictable group comprised single- and two-cue trials (rows 1, 2, 3, and 6). In each trial, the last cue validly indicated which location was probed at the end of the trial.

The cues were white central arrows, displayed for 100 ms, pointing to one of the six circle locations from the memory display. The last presented cues were always valid in predicting the location to be probed later. In case of two-cue trials, the first and the second cue never pointed to the same location. Finally, a probe color was presented in the original location of one of the six colored circles. The task was to compare the color of the circle in the probe display to the color of the circle at the same location from the memory display. The probe matched the memory display circle on half the trials (match trials). On the other half of trials the probe was either a color from another location in the memory display or a new color (mismatch trials). The probe remained on screen until the participant answered by pressing the left arrow key for a match and the right arrow key for a mismatch. Performance feedback

was immediately provided by a message on the screen (the German word for “right” or “wrong”). The feedback message disappeared after 500 ms and was followed by a gray inter-trial interval of 1.5 s. Speed of response was not emphasized. The predictability conditions were manipulated within-subject across sessions by providing only single-cue and no-cue trials in the predictable condition, and all cue types in the unpredictable condition (rows 1 to 5 in Figure 1).

To create conditions with matched retention intervals, and to prevent participants from anticipating the number of cues in a trial, the intervals between memory display and cue (memory-cue interval) as well as the intervals from the last cue to the probe (post-cue time) were manipulated. Our primary goal was to equate the overall retention interval, the time between the memory display and the valid cue, and the post-cue time for single-cue and two-cue trials; this was accomplished by the two conditions depicted in row 2 and row 3 of Figure 1. To prevent anticipation of the number of cues from the first memory display-cue interval, we also needed the conditions depicted in row 1 and row 4 in Figure 1. In Experiment 1A, we used a long and a short no-cue condition. The long interval no-cue condition was matched to the overall retention time of the three cue conditions and served as a baseline for evaluating the retro-cue benefit for the cue conditions at a constant retention interval. In the short interval no-cue condition, the onset of the probe was matched to the onset of the cue in the early-onset single-cue condition. The comparison of the early single-cue and the short no-cue condition tests whether there is still a retro-cue benefit if we assume that retrieval in the cue condition commences at the onset of the cue, whereas it starts at the onset of the probe in no-cue trials. In that case, the early-onset single-cue condition and the short no-cue condition are matched for their retention intervals.

Experiment 1B. Experiment 1B³ differed from Experiment 1A only regarding two aspects: First, predictability was manipulated between subjects by creating two experimental groups. This was done to rule out carry-over effects between the two predictability conditions. Second, we did not include no-cue trials. Instead, there was a two-cue late-onset condition, comparable to the two-cue condition in Experiment 1A, but the interval between memory display and first cue was 1600 ms (see row 6 in Figure 1). With the inclusion of this condition, we prevented anticipation of the number of cues even for trials with a late cue onset. The durations of all intervals for each condition can be seen in Figure 1. Single- and two-cue conditions matching with regarding to overall retention interval, valid cue onset time, and post-cue time are depicted in rows 2 and 3 of Figure 1. The conditions depicted in rows 1 and 6 of Figure 1 were needed to prevent the anticipation of the number of cues from the first memory display-cue interval.

Procedure

For Experiment 1A, each participant completed 12 practice trials prior to each test session. Each of the two test session consisted of 480 trials (eight blocks). Predictability conditions were manipulated across the two sessions with counterbalanced order. In the unpredictable condition, one third of trials were no-cue trials (equal number for long and short retention interval), one third were single-cue trials (equal number for early and late cue onset) and the remaining third were two-cue trials. In the predictable condition, half of the trials were no-cue trials (equal number for long and short retention interval) and the other half were single-cue trials (equal number for early and late cue onset). For Experiment 1B, each participant completed one session with 24 practice trials and 320 test trials (eight blocks).

³ Due to a programming error, red color probes appeared only in match, never in mismatch trials. Analyses with and without red probe trials converge. Analyses excluding red probe trials are reported.

There were equal numbers of trials from each condition in each experimental group; for the predictable group these were only the two kinds of single-cue trials, whereas for the unpredictable group these were single-trials and two-cue trials. Match and mismatch trials as well as the cue conditions were randomly intermixed in both experiments.

3.2.2 Results

For Experiment 1A, the average percentage correct across conditions was 77.2% (SD = 8.9). The average percentage correct in Experiment 1B was 81.4% (SD = 6.3) in the predictable and 81.5% (SD = 6.8) in the unpredictable group. All analyses for both experiments were conducted on the dependent measure percentage correct.

Retro-cue benefit

To test whether valid retro-cueing improves performance, we compared the no-cue condition (long) to the cue conditions; these conditions were matched regarding the overall retention interval in Experiment 1A. All comparisons revealed significant retro-cue benefits, showing that retro-cueing improved performance compared to not providing a cue. Performance in the single-cue (late) condition was significantly better than in the no-cue condition, both in the predictable session: $t(23) = 4.068$, $p < .001$, and in the unpredictable session: $t(23) = 6.304$, $p < .001$. Performance for single-cue trials (early) was also better than the no-cue condition, both in the predictable and in the unpredictable session, $t(23) = 9.031$, $p < .001$, and $t(23) = 6.173$, $p < .001$, respectively. Furthermore, performance in two-cue trials was significantly better than in no-cue trials (long), $t(23) = 3.848$, $p = .001$. Averages for each condition are shown in Figure 2, panel A.

Additionally, we compared the early single-cue condition to the no-cue condition (short) in Experiment 1A. This comparison provides a more conservative test of the retro-cue benefit because it equates the time between memory display and probe in the no-cue condition with the time between memory display and cue in the retro-cue condition (see rows 1 and 5 in

Figure 1). Thus, these two conditions are equated for the un-cued retention interval, that is, the time during which the entire memory set had to be retained in WM, without any information about which memory contents will be relevant at test. This comparison revealed a significant retro-cue benefit, for the predictable condition, $t(23) = 7.368$, $p < .001$, and for the unpredictable condition, $t(23) = 4.194$, $p < .001$.

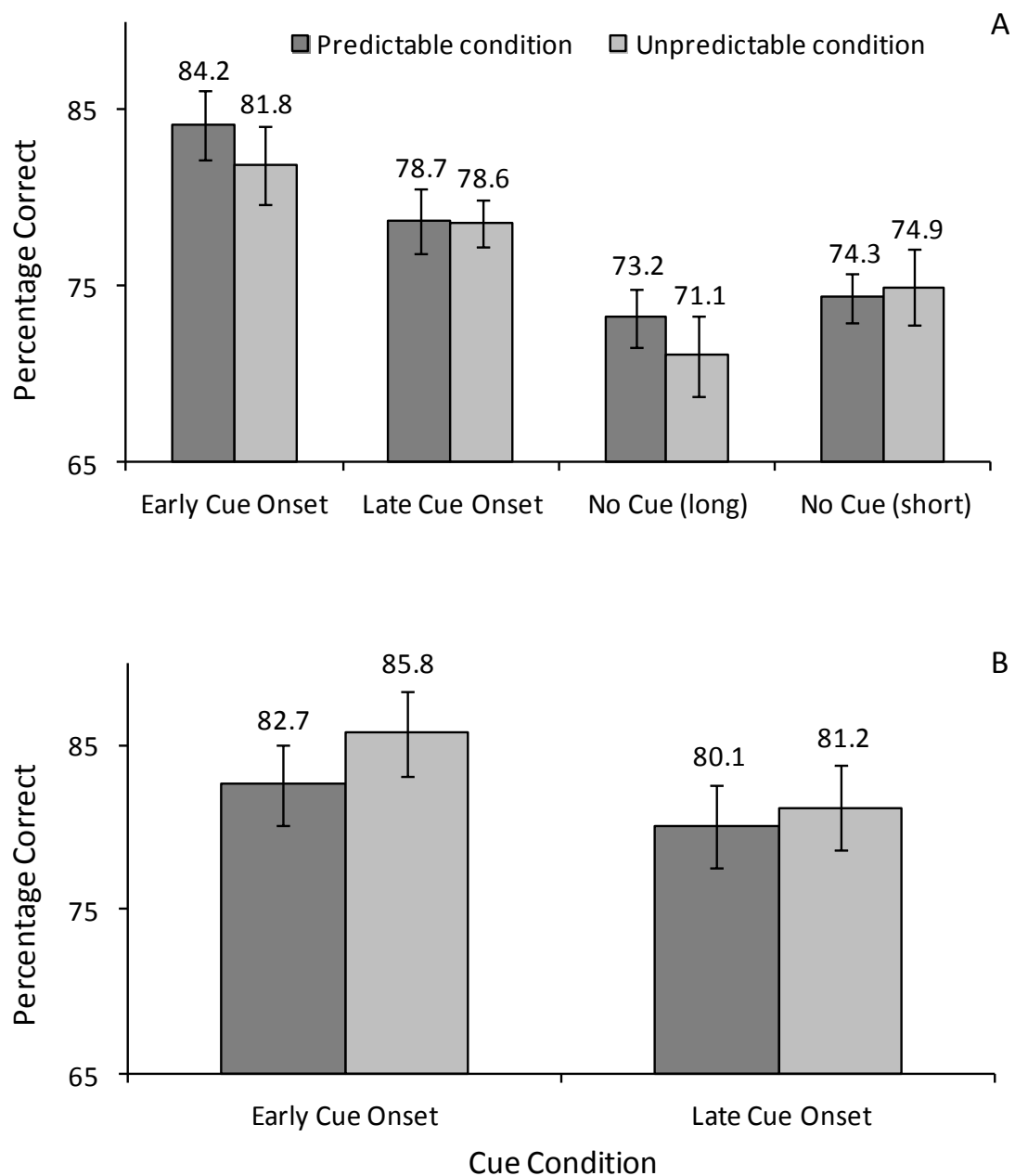


Figure 2. Panel A shows the percentage correct for single-cue and no-cue trials for Experiment 1A for the predictable and the unpredictable condition. Error bars represent within-subject confidence intervals (95%). Panel B shows the percentage correct for single-cue conditions for Experiment 1B for the predictable and the unpredictable group. Error bars represent between-subject confidence intervals (95%).

Predictability

To examine whether predictability influences WM performance, we conducted a within-subject ANOVA with the factors predictability (predictable and unpredictable condition) and cue onset (early or late) for single-cue trials for Experiment 1A (see Figure 2, panel A). Early cue onset led to better performance than late cue onset, $F(1, 23) = 26.967$, $p < .001$, partial $\eta^2 = .540$. The predictable and the unpredictable condition did not differ significantly from each other, $F(1, 23) = 1.633$, $p = .214$, partial $\eta^2 = .066$, and the interaction between the two variables was non-significant, $F(1, 23) = 2.144$, $p = .157$, partial $\eta^2 = .085$. Experiment 1B confirmed this pattern of results for the between-subject design: Performance was better for early single-cue onset than for late single-cue onset, $F(1, 58) = 35.818$, $p < .001$, partial $\eta^2 = .382$. There was no significant effect of predictability, $F(1, 58) = 1.557$, $p = .217$, partial $\eta^2 = .026$, and no significant interaction, $F(1, 58) = 2.717$, $p = .105$, partial $\eta^2 = .045$. Figure 2 (panel B) shows the averages for each experimental group (predictability) and cue onset condition.

Non-focused information in WM

To test whether non-cued information remains accessible in WM after cueing other information, a paired t-test was conducted, comparing performance in late single-cue trials to performance in two-cue trials for the unpredictable condition (rows 2 and 3 in Figure 1). These two conditions were matched with regard to the onset time of the valid cue, the post-cue time, and overall retention time. There was no statistically significant difference between the two. In Experiment 1A, mean accuracy in single-cue trial (late onset) was 78.6% (SD = 10.3) and the mean accuracy in two-cue trials was 77.2% (SD = 9.3), $t(23) = 1.018$, $p = .319$ (see Figure 3, panel A for means). Experiment 1B confirmed this pattern: Late single-cue trials and early two-cue trials in the unpredictable group did not differ significantly from each other (see Figure 3, panel B for means), $t(29) = .733$, $p = .469$. These findings imply that

while attention is focused on the first-cued item, the non-cued information remains unimpaired and can be focused on when a second cue follows.

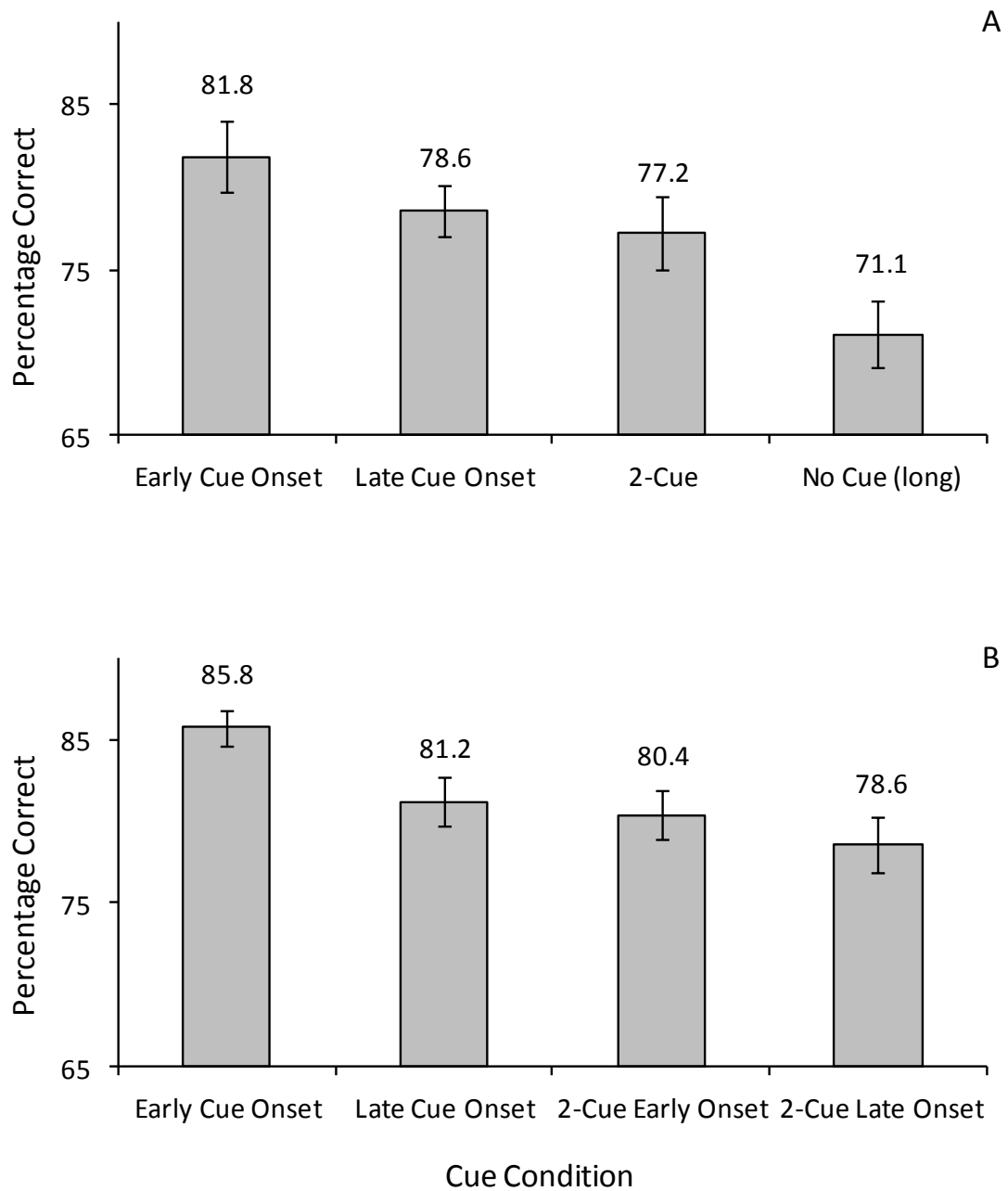


Figure 3. Panel A shows the percentage correct for single-cue trials (early and late onset), the two-cue condition, and the matched no cue condition in Experiment 1A (unpredictable condition). Panel B depicts accuracy for single-cue trials (early and late onset) and two-cue trials (early and late onset) in Experiment 1B (unpredictable condition). Error bars represent between-subject confidence intervals (95%).

3.2.3 Discussion

Experiment 1A showed that valid cue trials yield better performance than trials without any cueing, thus replicating the retro-cue benefit for our experimental procedure. This finding shows that our retro-cues were used to direct attention to WM representations.

Our finding of equivalent performance for single-cue and two-cue trials for matched retention intervals shows that focusing attention in WM on a single representation does not lead to forgetting of the other representations held in WM, supporting the maintenance hypothesis (see Table 1). This finding does not rule out forgetting over time, but it shows that if there is forgetting over time, it occurs independently of focusing.

This result is in line with the prediction from the three-embedded components model (Oberauer, 2002, 2009) and furthermore in agreement with the results of Landman et al. (2003), who did not find any significant differences between single-cue and two-cue trials in a similar experiment with a shorter time frame. This outcome indicates that the focus of attention can be flexibly re-oriented in WM and that focusing on one item does not compromise retention of the remaining items. However, the results are in conflict with the predictions from fixed-capacity and constant-resource theories, which predict that focusing one item reduces the number of slots or the amount of resources available for non-cued, non-focused WM contents.

Our result is in contrast to the one reported by Woodman and Vecera (2011), who found that performance declined over successive probes testing different items of a memory set. The reasons for this discrepancy could be that in our experiments people merely focused on several items in succession, whereas in the study of Woodman and Vecera, they were probed for overt responses on several items in succession. Probing memory for overt responses is known to generate output interference (Cowan, Saults, Elliot, & Moreno, 2002; Oberauer, 2003). Our finding that memory survives two successive retro-cues with no

measurable impairment suggests that merely focusing on an item does not create interference in the way that retrieving an item for an overt response does.

We found no effect of predictability of the number of cues, neither in the within- nor in the between-subject design. In single-cue trials, participants in the predictable session or group could have reduced WM load by removing the non-cued items from the central component of WM; in the unpredictable session or group, this strategy would have been detrimental. Nevertheless, performance in the two conditions was indistinguishable for single-cue trials. Together with the finding of equivalent performance in single- and two-cue trials, this implies that non-cued information was not removed but maintained in the predictable condition (for an overview of predictions, see Table 1), even when it was known to be irrelevant (cf. Makovski & Jiang, 2008).

It is conceivable that voluntary removal (as an indicator of flexible control) is a slow process that needs more time than was available to our participants. However, experiments on removal of information from verbal WM (Oberauer, 2001, 2002) indicate that it takes little more than 1 s to remove an irrelevant set from the central component of WM. In the early cue-onset conditions for single-cue trials participants had more than 1 s between the cue and the probe to remove irrelevant working-memory contents. This should have been sufficient time to allow for removal, and yet no evidence for removal was obtained. Hence, our finding suggests that non-cued information remains in WM, at least for the duration of our post-cue time.

Within the framework of null-hypothesis testing, any support for the null-hypothesis is necessarily indirect, stemming from a failure to support the alternative hypothesis. A direct assessment of the strength of evidence for the null hypothesis can be derived from likelihood ratios by estimating how much more likely it is to obtain the observed data under the assumptions of the null hypothesis compared to an alternative hypothesis (Glover & Dixon, 2004). In addition, the strength of evidence for the null hypothesis can also be assessed by the

Bayes Factor. From the Bayes Factor we can calculate the posterior probability that the null hypothesis is true, given the data, under the assumption of equal prior probabilities for the null hypothesis and the alternative hypothesis (Masson, 2011). In Table 2, we provide likelihood ratios, Bayes Factors, and posterior probabilities of the null hypothesis for the comparison of predictability conditions and the comparison of single-cue and two-cue trial performance, separately for Experiment 1A and 1B. In all cases, the data consistently support the null hypothesis over the alternative hypothesis. In the case of the predictability comparison, the support for the null hypothesis is actually stronger than the likelihood ratios and Bayes Factors suggest, because in the two experiments the deviation from the null hypothesis went in different directions—thus, we are testing the null-hypothesis against two different, mutually contradictory alternative hypotheses in Experiments 1A and 1B.

Table 2

Likelihood ratios in favor of the null hypotheses for Experiments 1A and 1B

Comparison	n	k1(k2)	LR _{BIC}	BF	P _{BIC} (H0 D)
E1A					
Predictable vs. unpredictable condition	24	2(3)	2.1505	2.1591	.6835
Single- vs. two-cue trials	24	2(3)	2.8874	2.8910	.7430
E1B					
Predictable vs. unpredictable group	60	2(3)	3.5470	3.5143	.7785
Single- vs. two-cue trials	30	2(3)	4.1584	4.1709	.8066

Note. n = number of participants; $k1$ = number of free parameters for null hypothesis (H_0); $k2$ = number of free parameters for alternative hypothesis (H_1); LR_{BIC} = Likelihood ratio in favor of the null hypothesis, corrected for number of free parameters according to BIC; BF = Bayes Factor (according to Masson, 2011); P_{BIC} ($H_0|D$) = posterior probability for the null hypothesis (Masson, 2011).

Our results suggest that WM representations can be maintained without continuous focal attention. This finding is in accordance with the three-embedded components framework of WM (Oberauer, 2002, 2009). It is also in accordance with several computational models of WM, such as the feature model by Nairne (1990), the primacy model by Page and Norris

(1998) and the Serial Order in a Box (SOB) model by Farrell and Lewandowsky (2002; Oberauer & Lewandowsky, 2008; Oberauer et al., 2012), all of which assume that maintenance in WM does not require focal attention.

3.3 Experiment 2: De-Focused Information in WM - Intrusion Approach

The first experiments showed that information that has never been focused after initial encoding is not impaired by focusing other information in WM. However, the case might be different for de-focused items in WM. These items gained a special status once by being focused during the retention interval, but were then de-prioritized again while another item is prioritized. We consider three hypotheses about the fate of de-focused items: the *back-to-baseline* hypothesis, the *refreshing* hypothesis, and the *inhibition* hypothesis. All of them are motivated by prior theorizing or data.

The back-to-baseline hypothesis is that de-focused items are set back to their status before being focused. Thus, their memory strength would not differ from that of other, never cued items. Bays and Husain (2008) investigated accuracy in a visual WM task as a function of the history of saccades to the locations of memory items during the retention interval. They observed a benefit for the last fixated item compared to previously fixated objects, but no advantage for any object fixated before the last. If fixating is an indicator of attentional focusing, this result would suggest that previously-focused objects return to baseline level, in line with the back-to-baseline hypothesis.

The refreshing hypothesis states that focusing on an item in WM could leave that item strengthened after the focus moved away. This possibility is implied by the notion of attention-based refreshing (Johnson et al., 2005; Raye, Johnson, Mitchell, Greene, & Johnson, 2007). Refreshing is assumed to be a mechanism strengthening an already active representation in memory by briefly thinking of it. Importantly, this happens in WM without perceptual input (Johnson, 1992; Johnson & Hirst, 1993). Hence, according to the refreshing

hypothesis, by focusing a representation in WM (i.e., refreshing it), its activation and presumably its binding to its context (e.g., its spatial location) are strengthened. In the context of a recognition test as in our experiments, it makes a difference whether refreshing only increases an item's activation or additionally strengthens bindings to its context, as we will explain below. In the data of Bays and Husain (2008), though the effect was not significant, there was a trend towards better accuracy for the next-to last fixated object compared to other objects. Possibly a lack of power prevented this trend, which would be predicted from the refreshing hypothesis, from becoming significant.

The inhibition hypothesis states that de-focused information is suppressed or removed from WM to facilitate disengagement of the focus of attention. Several researchers have assumed that inhibition is important to overcome prepotent responses and to prevent getting stuck in a task (Koch, Gade, Philipp, & Schuch, 2010; Mayr & Keele, 2000) or on an item (Klein, 2000; Maxcey-Richard & Hollingworth, 2012; Pratt, Kingstone, & Khoe, 1997). Phenomena such as inhibition of return (Klein, 2000) and response suppression (Hübner & Druey, 2006; Henson, 1998) suggest that leaving a once-selected representation behind is often accompanied by the suppression of that representation. Bao, Li, Chen, and Zhang (2006) have proposed that inhibition of de-focused items serves to facilitate focus switching in WM. In case de-focused information is inhibited, its representation is less accessible after de-focusing than representations of non-focused information.

We used the same basic paradigm as in Experiments 1A and 1B with some variations. The probe stimulus was presented centrally and had to be compared to the last cued object. The probe either matched or mismatched the cued item. We distinguish different classes of mismatching probes: New probes are objects that have not been presented in the memory display. Intrusion probes are objects that were included in the memory display, but in a different location than the one last cued. Accordingly, they require a rejection. Intrusion probes can be either created by presenting a not-cued object or the previously cued (i.e., de-

focused) object. We investigate whether rejecting previously-cued probes is harder or easier than rejecting new probes or non-cued intrusion probes, and infer from this whether de-focused information is maintained, strengthened, or inhibited in WM.

To make predictions for the three possible fates of de-focused information we need to consider the processes involved in recognition. According to the dual-process theory of recognition in WM (e.g., McElree & Doshier, 1989; Oberauer, 2008), two sources of information underlie recognition decisions: familiarity and recollection. Familiarity is assumed to reflect whether a stimulus has been encountered recently without considering its bindings to a spatial location or to other contextual information. Accrual of an internal familiarity signal in response to a probe is assumed to be automatic and fast. In the present paradigm, an assessment of familiarity would be sufficient for the rejection of new item probes. These probes do not elicit a strong familiarity signal and can be rejected quickly on that basis. In contrast, all items from the memory display have been encountered recently and elicit a strong familiarity signal, which is ambiguous because it does not discriminate between matching probes (to be accepted) and intrusion probes (to be rejected). To distinguish between matching probes and intrusion probes, recollection is required, which delivers information about the context of the familiar item in the memory display, such as its spatial location (Oberauer, 2008; Oberauer & Lange, 2009). This process is more time-consuming than familiarity assessment, but it enables a decision whether the probe matches the memorized object in the cued location.

On the basis of these considerations we make the following predictions for the three hypothetical consequences of de-focusing: If the de-focused item was simply maintained in a state equal to that before being focused, responses to previously-cued intrusion probes should not differ from responses to non-cued intrusion probes. If the de-focused item was suppressed in WM, rejecting a probe matching that item should be easier than rejecting an intrusion probe matching one of the non-cued items, because the suppressed items would elicit a weaker

familiarity signal, and this should be sufficient for rejecting the probes matching de-focused items. As a result, correct rejection rate of de-focused items should be as high as, or even higher than (in case of inhibition below baseline) those for new probes, and also higher than the correct-rejection rate for non-cued intrusion probes. Because the familiarity signal is available quickly, RTs for correct rejections should also be at least as fast for de-focused as for new probes, and faster than for non-cued intrusion probes.

In case de-focused information is strengthened in WM, two possibilities can be distinguished: One is that de-focused items remain highly activated in WM, without strengthening the bindings to their spatial locations. Their high activation would lead to strong familiarity, resulting in an increase of false alarms to these probe types, and slow RTs for correct rejections. The alternative possibility is that not only the activation but also the bindings of the de-focused item are strengthened, or that this item is bound to a representation of the fact that it had been de-focused before. Recollection of this binding information would facilitate “recall to reject” (Rotello & Heit, 1999): The de-focused item can be recollected together with its context and thereby be identified as not being the item in the currently cued location. This would enable participants to correctly reject de-focused intrusion probes with higher accuracy than non-cued intrusion probes. However, because recall-to-reject relies on recollection, these responses are still predicted to be slower than rejection of new probes.

In addition to the manipulation of mismatching probe types, we varied the time interval after the second cue (post-cue time, see Figure 4), to investigate whether it takes time to fully use the cue.

3.3.1 Method

Participants

Thirty-seven students participated in this experiment. Their mean age was 22 years (range: 19-28), and nine of them were male. Participants completed two sessions scheduled on

different days and received financial compensation (30 CHF for two one-hour sessions) or partial course credit in exchange for their participation. One participant had to be excluded from analysis due to technical problems.

Task and stimuli

The same color-recognition task as in Experiment 1 was used with some variation. Only single- and two-cue trials were provided and instead of presenting a location-bound probe (as in Experiments 1), the probe was presented centrally and had to be compared to the color of the last cued object. Thereby, the importance of processing the cue was stressed. A correct response was only possible if the cue was used to identify the relevant item. Of special interest were mismatch trials. These were created by different probe types: The probe was either a new color (not presented in the memory display), a non-cued color (either a spatial neighbor or a non-neighbor to the last cued location), or the color of the previously cued circle. Fifty percent of trials were positive probes (matches). The remaining mismatch trials were split into twenty-five percent new color probes and twenty-five percent intrusions. For single-cue trials, half of the intrusion trials were spatial neighbors and half non-neighbors to the last cued item. For two-cue trials, the intrusions were split into 12.5% previously cued item probes and 12.5% non-cued probes. Of the non-cued probes, half were spatial neighbors of the last cued item.⁴

Additionally, the time interval between memory display and first cue was kept constant and only the interval after the last cue (post-cue time) was varied: A short interval of 1100 ms for single-cue trials and 500 ms for two-cue trials, and a long interval of 1600 ms for single-cue trials and 1100 ms for two-cue trials (see Figure 4). This was done to test whether it takes time to make full use of the retro-cues. The short and long post-cue intervals differed

⁴ Due to a programming error, participants had on average 2.4 previously-cued probe trials more and 2.4 neighbor item probe trials less in condition 3 and 4 than originally intended.

between single-cue and two-cue trials because we held overall retention interval constant between these two kinds of trials.

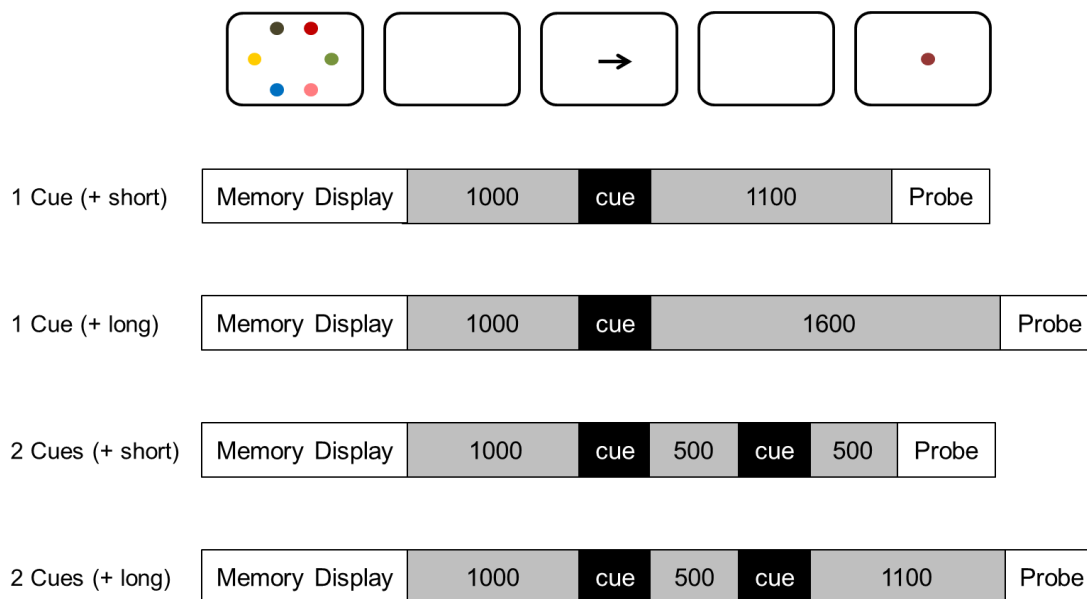


Figure 4. Overview of the experimental design and conditions for Experiment 2. The first two rows show the two single-cue trial conditions (with short and long post-cue time) and the third and fourth row show the two two-cue trial conditions (short and long post-cue time).

Procedure

Each participant completed 34 practice trials prior to the first test session and 4 practice trials prior to the second session. Each test sessions included 416 trials, split into nine blocks (the first eight including 50 trials each and the last one 16 trials). There was an equal number of trials for the four cue conditions. Cue conditions and probe types were randomly intermixed. Participants were instructed to respond as accurately and as fast as possible.

3.3.2 Results

Participants completed on average 82.5% of trials correctly ($SD = 6.4$). Percentage correct and RT data from the different categories of mismatch trials served as dependent variables. The data from the match trials are included in the analysis of post-cue time, but are omitted for analysis of the probe types. The mean scores and the corresponding standard deviations for each condition can be found in the Appendix.

Mismatch trials were sorted into 3 (single-cue trials) or 4 (two-cue trials) probe type categories: New, non-neighbor intrusion, neighbor intrusion, and previously cued intrusion probes (the latter only for two-cue trials). Trials associated with responses faster than 100 ms and longer than 7 seconds were excluded for the RT analyses, as were error trials. RTs were log-transformed for all analyses to reduce the skew of the distributions. To facilitate readability, un-transformed RT means are plotted in the figures and reported in the text.⁵ For some analyses of variance, the sphericity assumption was violated. In these cases, corrected Greenhouse-Geisser degrees of freedom (recognizable by non-integer values) were reported. For a graphical depiction of results (including means), see Figure 5.

Single-cue trials

A three-way ANOVA on probe type (new, non-neighbor intrusion, and neighbor intrusion probe) was run on accuracy data (percentage correct) for single-cue trials. The main effect was significant, $F(2, 70) = 11.98, p < .001$, partial $\eta^2 = .255$. Planned contrasts revealed that new-item probes did not differ significantly from intrusion probes (i.e., neighbor and non-neighbor probes combined), $F(1, 35) = 2.701, p = .109$, partial $\eta^2 = .072$. Neighbor probes led to significantly worse performance than non-neighbor probes, $F(1, 35) = 21.005, p < .001$, partial $\eta^2 = .375$.

The same ANOVA was run on log-transformed RT data. The main effect again was significant, $F(2, 70) = 3.375, p = .040$, partial $\eta^2 = .088$, reflecting faster RTs for new probes compared to the two kinds of intrusion probes combined, $F(1, 35) = 5.674, p = .023$, partial $\eta^2 = .139$. Neighbor and non-neighbor probes did not differ significantly from each other, as shown by planned contrasts, $F(1, 35) = 2.026, p = .163$, partial $\eta^2 = .055$. For a graphical overview, see Figure 5.

⁵ The same trimming procedure was used as for the log-transformed RTs. Additionally, RTs exceeding the participant's mean per design cell by more than three standard deviations were excluded.

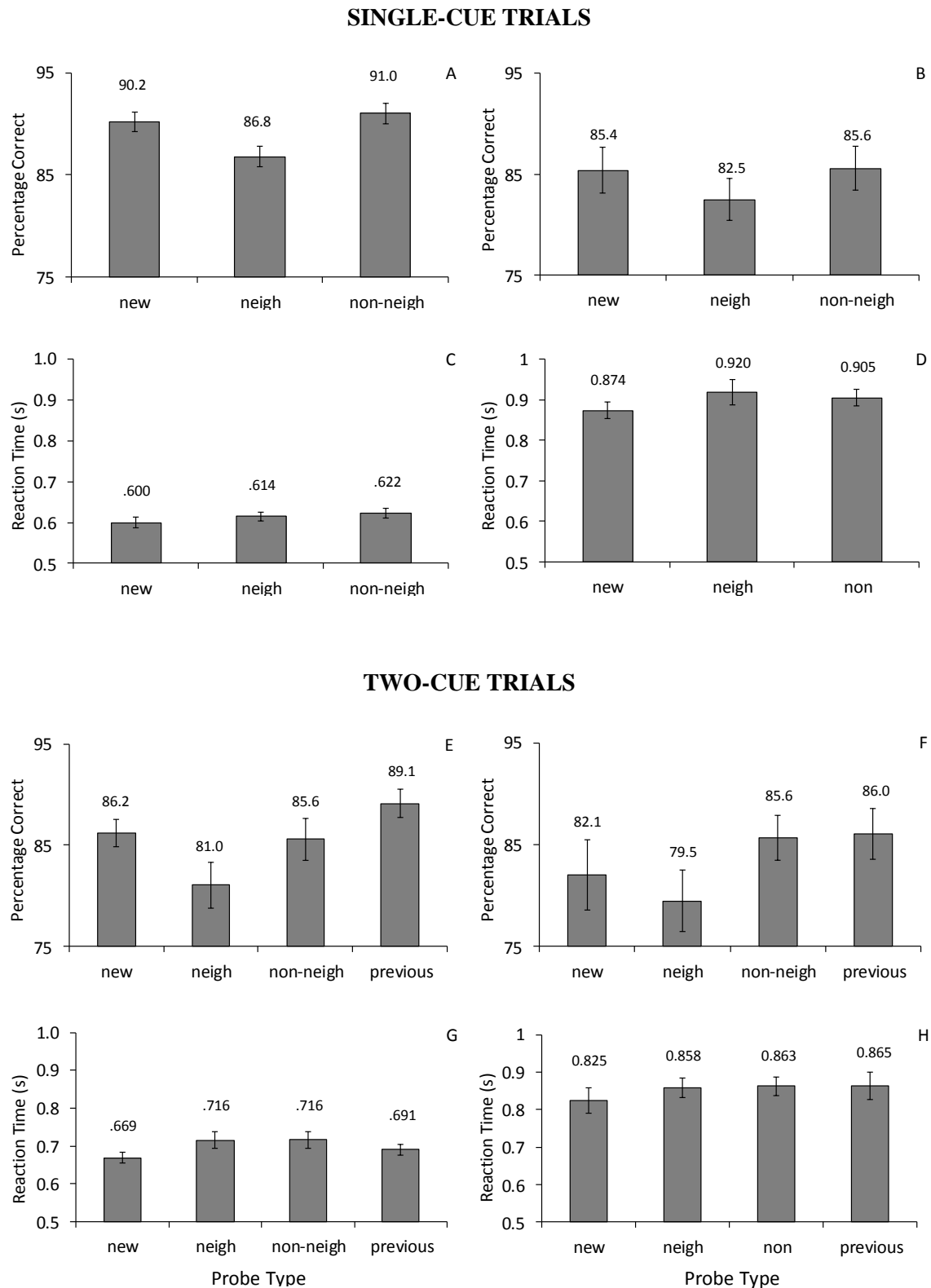


Figure 5. Results for Experiment 2 (first column) and Experiment 3 (second column). The first row shows accuracy data for Experiment 2 (panel A) and Experiment 3 (panel B) and the second row shows RT data for Experiment 2 (panel C) and Experiment 3 (panel D) for single-cue trials. The third row shows accuracy data for Experiment 2 (panel E) and Experiment 3 (panel F), and the fourth row shows RT data for Experiment 2 (panel G) and Experiment 3 (panel H) for two-cue trials. In each graph, the dependent variables are presented for each of the mismatch probe conditions (levels on the x-axis). Error bars represent within-subject confidence intervals (95%).

Two-cue trials

In two-cue trials, mismatches can be created by new color, non-neighbor color, neighbor color and previously cued color probes. A four-way ANOVA on probe type was conducted on accuracy data and on log-transformed RTs.

The analysis for the accuracy data showed that the main effect was significant, $F(2.085, 72.969) = 13.041$, $p < .001$, partial $\eta^2 = .271$. Planned contrasts revealed that new item probes did not differ significantly from all intrusion probes combined, $F(1, 35) = 1.466$, $p = .234$, partial $\eta^2 = .040$. Previously cued intrusion probes yielded significantly higher correct-rejection rates than neighbor and non-neighbor intrusion probes combined, $F(1, 35) = 37.304$, $p < .001$, partial $\eta^2 = .516$.⁶ Accuracy for neighbor probes was poorest, significantly worse than for non-neighbor probes, $F(1, 35) = 6.774$, $p = .013$, partial $\eta^2 = .162$.

The same analysis for log-transformed was not significant, $F(2.346, 82.110) = 1.055$, $p = .361$, partial $\eta^2 = .029$. Nevertheless, planned contrasts revealed that new color probes led to significantly shorter RTs than the three kinds of intrusion probes combined, $F(1, 35) = 5.765$, $p = .022$, partial $\eta^2 = .141$. Previously cued intrusion probes did not differ from neighbor and non-neighbor probes combined, which in turn did not differ significantly from each other, both $F_s < .004$, $p_s > .95$, partial $\eta^2 < .001$. For a graphical overview, see Figure 5.

Post-cue time

To examine the effect of post-cue time length, paired t-tests were conducted on this variable, separately for single- and two-cue trials as well as for the dependent measures accuracy and RT. The relevant means and test statistics are presented in Table 3. The analyses on accuracy showed no effect of post-cue time. In contrast, the RT analyses showed that participants answered significantly faster after a longer post-cue time.

⁶ Previously cued probes were themselves neighbors and non-neighbors of the validly cued probe with a ratio of approximately 2:3. The categories of neighbors and non-neighbors excluded previously cued probes.

Table 3

Mean RTs (in seconds) and Accuracies (in percent correct) in Experiment 2 as a function of post-cue interval

	Short Post-Cue Time	Long Post-Cue Time	<i>t</i> (35), <i>p</i>
RT, single-cue trials	0.59 (0.15)	0.58 (0.14)	3.36, <i>p</i> = .002
RT, two-cue trials	0.70 (0.20)	0.61 (0.17)	13.94, <i>p</i> < .001
Accuracy, single-cue trials	84.4 (6.6)	84.4 (6.6)	0.05, <i>p</i> = .96
Accuracy, two-cue trials	80.3 (7.0)	80.9 (7.0)	0.90, <i>p</i> = .33

Note. Standard deviations are depicted in parentheses.

3.3.3 Discussion

Experiment 2 aimed to answer the question of what happens to de-focused representations. The results showed that previously-cued item probes were rejected better than other intrusion probe types. They were responded to about as fast as other intrusion probes, and slower than new probes.

This pattern of effects can be explained within the dual-process theory of short-term recognition, distinguishing familiarity and recollection (McElree & Doshier, 1989; Oberauer, 2008). New probes elicit a weak familiarity signal, indicating that the object has not been encountered recently. This weak familiarity signal offers a shortcut: Because the probe appears unfamiliar, it was most likely not the validly cued representation. Hence, new probes can be rejected on the basis of a familiarity process, which is assumed to be fast (Göthe & Oberauer, 2008; Yonelinas, 2002). Intrusion probes elicit a stronger familiarity signal (due to having been encoded into WM recently) and therefore require a more time-consuming recollection process that retrieves information about which color was where in the memory display to reach a correct rejection. In line with the predictions of the dual-process theory,

intrusion probes - and among them the previously cued probes - exhibited slower RTs than new item probes.

The observation that previously cued probes were processed slowly, but accuracy on them was better than on the other mismatch probe types, suggests recall-to-reject: The previously cued item is maintained in WM after it has been de-focused, and the binding to its position remains strengthened. When the probe matches the previously cued item, it elicits a strong familiarity signal, which prevents fast rejection on the basis of low familiarity alone. Recollection reveals that the probe was part of the memory display but not in the relevant (i.e., last-cued) position. This information is available with high accuracy because of the strengthened item-position binding due to having been focused before. Therefore, previously-focused probes can be rejected with high accuracy.

The results regarding the post-cue time manipulation showed that the additional time after the last cue offset increased response speed for all mismatch-probe types in two-cue trials, although post-cue time had no effect on accuracy. This result provides tentative evidence that it takes time to make fully efficient use of the retro-cue.

Near probes were harder to reject than non-neighboring intrusion probes. This finding suggests that representations in WM are spatially imprecise (cf. Makovski & Jiang, 2008), such that when one tries to focus on an item in a given location, neighboring items cannot be completely excluded. A complementary pattern of results was obtained by Schmidt, Vogel, Woodman, and Luck (2002), using invalid location cues: They cued an item's location and probed the cued location, a neighboring location, or a non-neighboring location. Performance for neighboring locations was better than for non-neighboring locations, as would be expected if the effect of cueing spilled over to neighbors of the cued item.

The notion of a spatially imprecise WM assumes that representations of spatially separated objects in WM are not perfectly distinct. Rather, each object is bound to its location in space, and the location representations overlap as a function of their proximity. When one

location is cued, the location is used as a retrieval cue to the object bound to it. Because of location overlap, neighboring objects might be partially retrieved into the focus of attention, thereby causing interference from spatially close objects. A probe matching a neighbor of the cued item therefore matches part of the information in the focus, biasing the decision for neighbor probes towards a match response, although a rejection is required. The notion of imprecise, overlapping spatial representations in WM for simultaneous multi-object displays is analogous to the well-established notion of overlapping temporal locations in WM for sequentially presented items (Brown, Preece, & Hulme, 2000; Burgess & Hitch, 1999; Lewandowsky & Farrell, 2008).

In sum, the results provide initial support to the refreshing hypothesis, according to which focusing an item strengthens its binding to its spatial location, thereby enabling recall-to-reject. However, an alternative explanation for the high accuracy on probes matching defocused items could be based on the inhibition hypothesis, assuming that the previously cued item is inhibited to prevent interference. This in turn should make rejection of a previously-cued item easier because a probe matching an inhibited item generates less familiarity than a probe matching a non-cued item. The fact that previously cued items yielded rather slow RTs does not sit comfortably with the inhibition hypothesis, because a probe with low familiarity should be rejected quickly, as was the case for new probes. To be made compatible with the inhibition hypothesis, the RT data pattern would have to be explained by a speed-accuracy trade-off acting specifically on previously-cued probes. This explanation appears unlikely, because it requires a paradoxical assumption: The presumed speed-accuracy trade-off would have to occur selectively for previously-cued probes. Thus, the cognitive system would first have to identify a probe as matching a previously cued item, to then shift the speed-accuracy criterion for the process of identifying whether or not the probe matches the last-cued item. The absurdity of this notion renders an inhibition hypothesis of the present data implausible. Nevertheless, because of the remaining ambiguity of how to interpret the high rejection

accuracy for previously-cued probes, we carried out a further experiment to explicitly test whether de-focused items are inhibited or strengthened.

3.4 Experiment 3: De-Focused Information in WM – Backshift Approach

Experiment 3 serves to decide between the refreshing and the inhibition hypothesis. We presented up to three cues in succession. This enabled us to investigate sequences in which the focus of attention was first cued to one item, then away from it, and finally back to it (sequence ABA), in comparison to three-cue trials in which three different items were cued (sequence CBA). In both trials, the last-cued item A is indicated as relevant for the comparison with the probe. In the ABA sequence, that item has been previously cued, whereas in the CBA sequence it has not. If focusing on an item strengthens its representation in WM, then the comparison of the probe should be more successful in trials with cueing sequence ABA than in trials with CBA sequence. In contrast, if a previously cued item is inhibited, comparison of the probe to the finally cued item (A) should be impaired in ABA sequences relative to CBA sequences (see Mayr & Keele, 2000, for a demonstration of inhibition of previously-used task sets based on this rationale).

3.4.1 Method

Participants

Twenty-seven students participated in the experiment. Their mean age was 26 years (range 21 to 35), and nine of them were male. Participants received financial compensation or course credit. One participant had to be excluded due to accuracy below the predefined exclusion criterion (65% correct).

Task and stimuli

The same basic paradigm was used as in Experiment 2 with some adaptations (see Figure 6 for an overview of the flow of events in each condition, including interval durations).

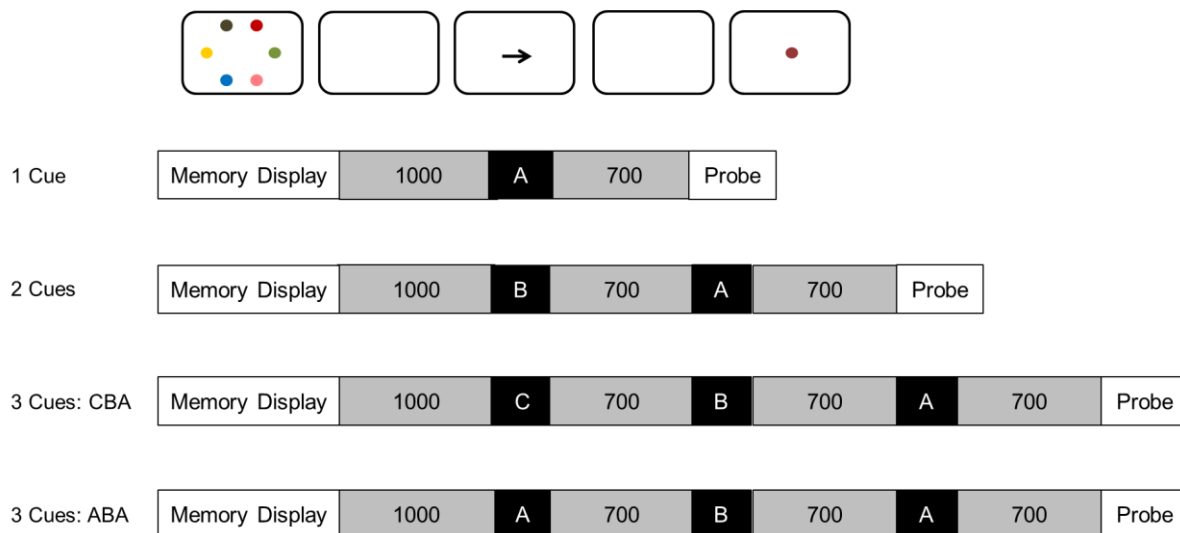


Figure 6. Overview of the experimental design and conditions for Experiment 3. The first row depicts a single-cue trial, the second row a two-cue trial. In the two last rows, the two different three-cue trial sequences are shown: The third row shows the CBA and the fourth the ABA sequence. The numbers in the frames depict the duration of the corresponding intervals in milliseconds.

In the interval between memory display and probe display, either one, two, or three cues were presented sequentially with an inter-cue-interval of 700 milliseconds. Participants' task was to compare the central probe item to the item which was cued last in the trial. The probe item could be either a match or a mismatch. Mismatches were classified into new color probes, neighbor, non-neighbor, and previously-cued probe items. In three-cue trials, we did not distinguish neighbors and non-neighbors but rather collapsed them into one category – non-cued item probes. Three-cue trials consisted of CBA and ABA sequences. The CBA sequences cued three different locations selected at random without replacement, whereas in the ABA sequence the first and the third cue pointed to the same location (the second cue was selected at random from all items except A). The number of trials per probe condition can be found in Table 4.

Table 4

Overview of the number of trials per probe type in Experiment 3

Cue Condition	Probe Type			
	Match	New	Non-Cued Intrusion	Previously Cued Intrusion
Single Cue	120	40	40 (non), 40 (neigh)	---
Two Cues	120	30	30 (non), 30 (neigh)	30
Three Cues				
CBA	192	48	96	48
ABA	48	12	24	12

Note. Non-cued intrusions = probes matching an object that was not cued in the corresponding memory display; non = non-neighbors; probes matching an object that was not a neighbor of the last cued object; neigh = neighbors; probes matching a neighbor of the last cued object.

Procedure

Each participant completed 24 practice trials prior to each of the two test sessions. Each session lasted about one hour and included 480 test trials, split into 10 blocks (the first nine comprising 50, the last one 30 trials). The different cue conditions and probe types were presented in random order. Participants were instructed to answer as fast as possible, without committing errors.

3.4.2 Results

Percentage correct as well as RT data served as dependent variables, and data trimming and transformation was conducted as for Experiment 2. The overall average percentage correct was 78.8% (SD = 7.1).

Number of cues

Our first analysis investigates whether information in WM can be maintained across up to three successive cues. A repeated-measures ANOVA on overall percentage correct for the three cue conditions (single-cue: mean = 80.1%, SD = 7.9; two-cue: mean = 77.9%, SD =

8.0; and three-cue: mean = 78.6%, SD = 7.2) was marginally non-significant, $F(2, 50) = 2.963$, $p = .061$, partial $\eta^2 = .106$. Pairwise comparisons revealed significantly better performance for single-cue than for two-cue trials ($p = .015$). The comparisons of single-cue trials to three-cue trials, as well as the comparison of two-cue trials to three-cue trials yielded non-significant outcomes ($p = .161$, and $p = .421$, respectively). The same repeated-measures ANOVA on log-transformed RT was significant, $F(2, 50) = 66.384$, $p < .001$, partial $\eta^2 = .726$. Pairwise comparisons revealed significant differences for each single comparison (all p -values $< .001$), showing that single-cue trials (untransformed mean = .869 s, SD = .223) led to the longest RTs and three-cue trial (untransformed mean = .727 s, SD = .175) to the shortest RTs, two-cue trials (untransformed mean = .819 s, SD = .208) ranging between them. The opposite trends on accuracy and response speed suggest that with increasing number of cues participants tended to trade accuracy for speed. Together, there was little evidence in the present results of memory loss over a larger number of successive cues. Apparently, people have little difficulty shifting their focus of attention successively to up to three different items in WM.

Three-cue trial sequences

To explore the main question of whether the previously cued item is strengthened in WM or whether it is inhibited, a paired t -test was conducted for repeated measures, comparing the two three-cue trial sequences (CBA and ABA) to each other. The analysis for percentage correct revealed better performance for the ABA sequence (mean = 82.5%, SD = 7.2) than for the CBA sequence (mean = 77.7%, SD = 7.4), $t(25) = 6.075$, $p < .001$. The same analysis on log-transformed RTs confirmed the pattern, with faster RTs for the ABA sequence (untransformed mean = 0.698 s, SD = 0.176) than for the CBA sequence (untransformed mean = 0.734 s, SD = 0.175), $t(25) = -7.239$, $p < .001$. Both results unambiguously support

the conclusion that the previously cued item is strengthened, rather than inhibited, upon leaving the focus of attention.

Furthermore, the ABA sequence produced significantly better performance than two-cue, $t(25) = 4.844$, $p < .001$, and single-cue trials, $t(25) = 2.988$, $p = .006$, and significantly faster RTs, $t(25) = 10.281$, $p < .001$ and $t(25) = 13.255$, $p < .001$, respectively. The CBA sequence led to accuracy indistinguishable from two-cue trials, $t(25) = .273$, $p = .787$, whereas RTs in the CBA trials were even faster than in the two-cue trials, $t(25) = 6.255$, $p < .001$.

This experiment also offers another opportunity to look at intrusion effects from previously cued item probes (previously cued item refers to the first cued item in two-cue trials and to the second cued item in three-cue trials).

Mismatch probe type pattern

Analyses on log-transformed RT and accuracy data across mismatch probe conditions were conducted separately for each cue condition (single-cue, two-cue, and three-cue trials). The results of the planned contrasts on mismatch probe type (new, neighbor intrusion, non-neighbors intrusion, and previously-cued intrusion probes) revealed the same overall pattern for each of the cue conditions, replicating the findings from Experiment 2. Results for each cue condition (single-cue, two-cue, and the CBA sequences) were computed separately. The ABA condition provided only few data per cell (see Table 4) and therefore, analysis of ABA data was omitted.

Accuracy on new probes did not differ from accuracy on all other mismatch probes combined. Previously-cued probes led to significantly higher accuracy than non-cued probes combined. In single-cue and two-cue trials, neighbor probes resulted in worse performance than non-neighbor probes (for the CBA sequence we did not distinguish neighbor and non-neighbor probes because of the sparseness of data). Regarding RTs, new item probes were responded to faster than all intrusion probes combined, and previously cued item probes did

not differ from the non-cued probes combined. Furthermore, non-neighbors and neighbors did not differ significantly from each other in single-cue and two-cue trials. The test statistics can be found in Tables 5 and 6, and Figure 5 provides a graphical overview of the results. Averages and the corresponding standard deviations for each condition (RT and accuracy) can be found in the Appendix.

Table 5

Planned contrasts on accuracy data for the different mismatch trial conditions in Experiment 3

Contrast	F-Value	df1	df2	partial η^2	p
Single-cue trials					
New vs. [non-neighbors & neighbors]	.881	1	25	.034	.357
Non-neighbors vs. neighbors	4.423	1	25	.150	.046
Two-cue trials					
New vs. [previous & non-neighbors & neighbors]	.669	1	25	.026	.421
Previous vs. [non-neighbors & neighbors]	5.499	1	25	.180	.027
Non-neighbors vs. neighbors	11.001	1	25	.306	.003
Three-cue trials					
New vs. [previous & non-cued]	.094	1	25	.004	.761
Previous vs. non-cued	25.940	1	25	.509	<.001

Note. Planned contrasts for single-, two-, and three-cue trials (CBA sequences) for percentage correct.

Table 6

Planned contrasts on log-transformed RT data for mismatch trial conditions in Experiment 3

Contrast	F-Value	df1	df2	partial η^2	p
Single-cue trials log-RT					
New vs. [non-neighbors & neighbors]	14.639	1	25	.369	.001
Non-neighbors vs. neighbors	.009	1	25	.000	.927
Two-cue trials log-RT					
New vs. [previous & non-neighbors & neighbors]	5.530	1	25	.181	.027
Previous vs. [non-neighbors & neighbors]	.108	1	25	.004	.745
Non-neighbors vs. neighbors	.006	1	25	.000	.938
Three-cue trials log-RT					
New vs. [previous & non-cued]	21.445	1	25	.462	<.001
Previous vs. non-cued	.208	1	25	.008	.652

Note. Planned contrasts for single-, two-, and three-cue trials (CBA sequences) for log-transformed RT data.

3.4.3 Discussion

The comparison of the two three-cue trial sequences revealed significantly better performance and faster RTs for the ABA sequence than the CBA sequence. This finding rules out the hypothesis that previously focused items are inhibited when another item is focused. If the previously focused item was inhibited, a shift back to that item in response to the final cue would have resulted in worse, not better performance. The results of Experiment 3 further corroborate the conclusion that previously focused items remain strengthened, including strengthened item-to-context bindings. When these items are not cued as relevant by the final cue, their increased accessibility enables efficient recall to reject, as in Experiment 2. When these items are cued again as relevant by the final cue, as in the ABA sequences of the present experiment, their increased accessibility enables acceptance of matching probes with high speed and accuracy.

Our result is in line with an analogous observation in the multiple-probe study of Woodman and Vecera (2011). They probed participants on features of different objects in varying orders of objects: In their ABAC sequence, the third object tested was the same as the first; hence, this sequence requires a backshift similar to the one in our ABA sequence. In their ABCA sequence, the third object is different from all previously tested objects, corresponding to our CBA sequence. When comparing performance on the third probe between those two sequences, performance in the ABAC sequence was considerably better than performance for in the ABCA sequence. This matches the pattern we obtained by comparing the CBA to the ABA sequence, thereby supporting our finding of strengthened bindings due to focusing.

Replicating the findings from Experiment 2, previously cued probes showed equally slow RTs as non-cued intrusion probes. Hence, all intrusion probes from the memory display - including previously cued ones - resulted in significantly slower RTs than new item probes. This pattern is in agreement with the dual-process theory of familiarity and recollection in WM (Oberauer, 2008), which postulates fast rejection of new item probes on the basis of their relatively weak familiarity signal, and a slower recollection process to discriminate between probe types generating a stronger familiarity signal, such as intrusion probes and matching probes. The relatively slow response times for previously-cued probes, in combination with high accuracy of rejecting them, further bolsters the assumption that the fate of de-focused information in WM can be explained by the refreshing hypothesis: Temporarily focusing on an item strengthens its bindings to its location without reducing its activation, so that it remains (at least) equally familiar as non-cued items, but easier to recollect.

Finally, we replicated the finding from Experiment 2 that neighbor probes were most difficult to reject. This finding further supports the notion of spatially imprecise representations in WM.

3.5 General Discussion

The present study investigated the fate of non-focused information, and the fate of previously focused and then de-focused information in WM.

3.5.1 The Fate of Non-Focused Information in WM

The first two experiments (1A and 1B) revealed consistently that focusing one item in WM does not impair memory for the remaining non-focused items. When valid cue onset and post-cue time were kept constant, single-cue and two-cue trials were responded to with equal accuracy.

Our pattern of results is difficult to accommodate by the discrete-capacity and the constant-resource theories. Both theories predict that focusing on one representation in WM compromises the quality or accessibility of non-focused information. In the resource model, activation is shared among items in the central component of WM. This implies that giving more activation to a single representation – the one prioritized by the first cue – draws activation away from the remaining non-focused items. For two-cue trials this implies that when the second cue appears, it points to a representation with reduced activation, implying that this item can no longer be recovered with high accuracy. As a consequence, re-focusing in two-cue trials should lead to worse performance compared to single-cue trials. The same prediction derives from the discrete-capacity theory, which assumes that focusing benefits are obtained by assigning more than one slot to the cued representation. Accordingly, focusing one item comes at the expense of one non-focused item which must give up its slot. As a consequence, in two-cue trials the first cue increases the chance that the item cued second is no longer available because it has lost its slot to the first-cued item. This implies that, on average, accuracy on two-cue trials should be worse than on single-cue trials.

In contrast to this prediction, we obtained no difference between single-cue and two-cue trials. This result was predicted by the three-embedded components model (Oberauer,

2002, 2009): Due to being in the focus of attention, the cued representation is particularly accessible, while the other representations in the direct-access region remain unaffected.

A comparison between a predictable group that could have discarded all non-cued items upon seeing the first cue, and an unpredictable group that could not, failed to reveal a difference in accuracy in single-cue trials in both Experiment 1A and 1B. This result indicates that both groups maintained the entire memory display in WM, even though removal of the non-cued items would have been optimal for the predictable group. Apparently, in the designs used in these experiments, non-focused information is maintained whether needed or not. In the present experimental design, where one randomly selected item is cued as relevant, it seems to be difficult to remove all the other items from WM. In contrast, removal of irrelevant information is possible in other situations, as shown by evidence from a modified Sternberg task (Lewis-Peacock et al., 2012; Oberauer, 2001, 2005b). One potentially relevant difference between studies in which removal effects were obtained and those in which no removal was observed is the structure of WM contents.

In the present study, a single randomly selected item was cued at any one time, and after the last cue, all other presentations turned irrelevant. The irrelevant objects in this task did not form a pre-defined set. Rather, the set of irrelevant information had to be formed post-hoc, after the cue indicated the relevant item. In contrast, studies showing successful removal of irrelevant information did not cue single items but entire sets. In these tasks, the sets were pre-defined at encoding (Matsukura et al., 2007; Oberauer, 2001, 2002, 2005b): Two clearly distinguished lists or sets were initially encoded and remembered, and then one set was designated as relevant. Perhaps removal can be applied only to distinct sets of irrelevant information. Constructing these sets post-hoc might require additional time, which was limited by the interval between the last cue and the probe. Therefore, although the time we provided for removal might have been long enough to remove a predefined set – as suggested

by evidence from several studies (Oberauer, 2001, 2002, 2005b) – this time might not have been sufficient for constructing a set and removing it.

3.5.2 The Fate of De-Focused Information in WM

The second question we addressed concerns the fate of de-focused information in WM. Experiments 2 and 3 showed that focusing an item strengthens its bindings to its context, and this gain in strength remains when the item is de-focused later. Thereby, the de-focused item can be recollected with high accuracy, resulting in a relatively high correct rejection rate. The advantage for the ABA sequence over the CBA sequence in Experiment 3 further supports the refreshing hypothesis but contradicts the inhibition hypothesis.

None of the three theoretical frameworks considered in the introduction makes explicit predictions for the fate of de-focused information, and none of them seems to readily be capable of explaining our evidence for strengthening of de-focused information without making further assumptions. In the following paragraphs, we discuss the results in light of the three theories and propose assumptions that could make them account for strengthening of de-focused items.

The discrete-capacity theory can explain focusing by assuming that the focused item receives two slots. Extending this idea, we could assume that the de-focused item retains its two slots even after the second item was cued. However, this would seriously reduce the accessibility of the non-focused information. After two successive cues, four slots would be used, two for each of the two cued items. Assuming that the average young adult has four slots available (Zhang & Luck, 2008; Cowan, 2005), no slot would be left, on average, for the remaining non-cued items. Accordingly, we should expect performance hardly better than chance in the CBA sequence, in which three different items are cued, in contrast to much better performance in the two-cue (BA) sequence, in which only two items are cued. In fact, performance hardly differed between these two conditions.

The same argument applies to the resource model, although this model is more flexible because activation can be divided up among items in continuous quantities. Activation is assumed to be a limited resource, such that giving more activation to one representation comes at the expense of other representations' activation. Strengthening of de-focused items could be explained within this model by assuming that the de-focused item retains some of its extra activation after the next item was cued. This would imply that the more items are cued in succession, the less activation is left for the remaining items, including the one cued last, and probed. Although the resource model, by assuming that only a small quantity of extra activation remains with de-focused items, can avoid predicting catastrophic forgetting for the CBA sequence, it still must predict a measurable performance decline for the CBA sequence compared to the BA sequence. No such decline was observed.

A full explanation of our results concerning de-focused items must include the assumption that de-focused items are left behind not just with higher activation, but with stronger bindings to their location. This becomes obvious when comparing the effects of using a de-focused item probe to the effects of presenting a neighbor color as probe stimulus. The de-focused item is rejected better compared to the other intrusion probes combined (non-neighbors and neighbors) whereas the neighbor probe is rejected worse than all other mismatch probe types. Within the constant-resource model, the neighborhood effect can be explained by assuming that neighbors of the last-cued (and probed) item receive some of the extra activation assigned to the cued item, perhaps by a spatially imprecise allocation of activation. As a consequence, probes matching a neighbor elicit a stronger familiarity signal than probes matching non-neighbors. This leads to more false alarms to neighbors than to non-neighbors. In contrast, probes matching de-focused items do not lead to more, but to fewer false alarms. This contrast shows that the neighborhood effect and the effect of de-focusing cannot be explained by the same mechanism. Therefore, a resource theory cannot at the same time explain the neighborhood effect by assuming that activation spills over from

the focused item to its neighbors, and explain strengthening of the de-focused item by assuming that the de-focused item retains some of its extra activation.

According to the three-embedded components framework (Oberauer, 2002, 2009) items are held in the region of direct access by virtue of being bound to their contexts, and focused information is held in a separate focus of attention. The special status of the item selected as the content of the focus of attention explains the retro-cueing benefit. The theory so far makes no assumptions about de-focused items. By default, de-focused information would be expected to remain in the region of direct access unmodified. To account for our finding of strengthening of de-focused items, the model has to be augmented by the assumption that focusing an item strengthens its binding to its context in the direct-access region, and these bindings remain strengthened even when the focus moves away to another item.

3.5.3 Representations in WM are Spatially Imprecise

In addition to the answers to the main research questions outlined above, our experiments showed that representations in WM are spatially imprecise. Consistently, probes matching items spatially adjacent to the cued item (i.e., neighbor probes) led to worse accuracy than the other intrusion probes. This pattern could be explained by spatially non-discrete representations in WM: Objects bound to spatially neighboring locations are less distinct in memory than objects in more distant locations. When an object cued by its location is retrieved into the focus of attention, the lack of spatial distinctiveness implies that not only the cued object itself, but also some information from the neighboring objects enters the focus. When a neighboring item probe is then presented, it matches the partial information from this neighbor in the focus, resulting in a misleading match-signal that causes the frequent false alarms to neighboring item probes.

One implication of this view is that the focus of attention is not strictly limited to a single item – rather, the content of the focus can consist of a blend of several items. This idea matches well with our assumption that the focus of attention is not structurally limited. Rather, the focus of attention is a mechanism for selecting individual items in central WM (Oberauer & Hein, 2012). It is limited to a single item to the extent that this selection is successful. The more items are held in the direct-access region at the same time, and the more the contexts they are bound to overlap, the more difficult it is for the focus to narrow down on a single item at the exclusion of others.

3.6 Conclusion

In sum, the results of all experiments converge on the conclusion that WM representations that have never been focused remain accessible in the region of direct access. Focusing an item strengthens the bindings to its context, and these bindings remain strengthened in the region of direct access after the item has been de-focused. Whereas the maintenance of non-focused information can be readily explained by the three-embedded components model (Oberauer, 2009), two alternative theoretical frameworks, discrete-capacity theories and constant-resource theories, cannot easily explain this result. The observation of strengthened item-context binding of de-focused information was not predicted by either model. It can be accommodated by the three-embedded-components model with the additional assumption that focusing refreshes item-context bindings in the direct-access region. The alternative theories considered here face difficulties accommodating this result.

3.7 Appendix

Table A1

RT data per probe type and cue condition in Experiment 2

Cue Condition	Match	Probe Type			De-focused
		New	Non-Cued		
Single Cue short	.569 (.146)	.603 (.139)	non: .627 (.178); neigh: .631 (.183)		---
Single Cue long	.546 (.139)	.599 (.148)	non: .618 (.151); neigh: .597 (.139)		---
Two Cues short	.661 (.191)	.712 (.177)	non: .773 (.285); neigh: .759 (.287)		.747 (.223)
Two Cues long	.586 (.186)	.625 (.183)	non: .658 (.187); neigh: .672 (.167)		.633 (.166)

Note. Non-cued = intrusion probes that were not cued in the corresponding memory display. De-focused = de-focused intrusion probe. Non = non-neighbor probes, neigh = neighbor probes. RT data are given in seconds (non-transformed), trimmed as described in the Method section. Standard deviations are provided in parentheses.

Table A2

Accuracy data per probe type and cue condition in Experiment 2

Cue Condition	Match	Probe Type			De-focused
		New	Non-Cued		
Single Cue short	79.7 (10.2)	89.8 (5.8)	non: 91.3 (7.6); neigh: 85.6 (10.8)		---
Single Cue long	78.8 (10.6)	90.6 (5.7)	non: 90.7 (8.8); neigh: 87.9 (6.8)		---
Two Cues short	74.7 (10.3)	85.6 (7.8)	non: 87.2 (12.7); neigh: 79.3 (11.9)		88.3 (8.1)
Two Cues long	74.8 (11.6)	86.8 (5.9)	non: 84.0 (11.4); neigh: 82.9 (12.9)		90.0 (8.4)

Note. Non-cued = intrusion probes that were not cued in the corresponding memory display. De-focused = de-focused intrusion probe. Non = non-neighbor probes, neigh = neighbor probes. Standard deviations are given in parentheses.

Table A3

RT data per probe type and cue condition in Experiment 3

Cue Condition	Match	Probe Type			De-focused
		New	Non-Cued		
Single Cue	.835 (.238)	.874 (.216)	non: .905 (.225); neigh: .920 (.242)		---
Two Cues	.780 (.210)	.825 (.215)	non: .863 (.217); neigh: .858 (.214)		.865 (.258)
CBA	.696 (.176)	.735 (.170)	.784 (.199)		.779 (.165)
ABA	.647 (.176)	.732 (.172)	.733 (.187)		.773 (.237)

Note. Non-cued = intrusion probes that were not cued in the corresponding memory display. De-focused = de-focused intrusion probe. Non = non-neighbor probes, neigh = neighbor probes. RT data are given in seconds (non-transformed), trimmed as described in the Method section. Standard deviations are given in parentheses.

Table A4

Accuracy data per probe type and cue condition in Experiment 3

Cue Condition	Match	Probe Type			De-focused
		New	Non-Cued		
Single Cue	75.6 (11.7)	85.4 (9.9)	non: 85.6 (9.7); neigh: 82.5 (7.9)		---
Two Cues	72.5 (10.3)	82.1 (12.0)	non: 85.6 (10.8); neigh: 79.5 (11.3)		86.0 (11.2)
CBA	74.4 (9.7)	81.5 (9.3)	78.2 (8.7)		85.7 (8.8)
ABA	78.8 (9.4)	87.2 (12.1)	85.7 (11.8)		85.9 (11.7)

Note. Non-cued = intrusion probes that were not cued in the corresponding memory display. De-focused = de-focused intrusion probe. Non = non-neighbor probes, neigh = neighbor probes. Standard deviations are given in parentheses.

4 Spatial Transposition Gradients in Visual Working Memory

Laura Hein, Klaus Oberauer, & Hsuan-Yu Lin

University of Zurich

Submission status:

Submitted for publication

Authors' contributions:

Laura Hein: Literature review, development of the research question and goals, design of the study, programming experiments in MATLAB, data collection, analysis, and interpretation, writing and submission of the manuscript

Klaus Oberauer: Supervision and discussion Laura Hein's contributions and revision of the manuscript

Hsuan-Yu Lin: Data aggregation and simulation for Experiment 2A and 2B and assistance in the data analysis of these experiments

Acknowledgements:

This research was funded by the Swiss National Science Foundation (project number 100014_126766 / 1). We thank Matthias Hartmann and Iliana Karipidis for the data collection.

Abstract

In list memory, access to individual items reflects limits of temporal distinctiveness. This is reflected in the finding that neighboring list items tend to be confused most often. This article investigates the analogous effect of spatial proximity in a visual working-memory task. Items were presented in different locations varying in spatial distance. A retro-cue indicated the location of the item relevant for the subsequent memory test. In two recognition experiments, probes matching spatially close neighbors of the relevant item led to more false alarms than probes matching distant neighbors or non-neighboring memory items. In two probed-recall experiments, one with simultaneous, the other with sequential memory item presentation, items closer to the cued location were more frequently chosen for recall than more distant items. These results reflect a spatial transposition gradient analogous to the temporal transposition gradient in serial recall and challenge fixed-capacity models of visual working memory (WM).

Keywords: transposition errors, spatial distinctiveness, cue-based retrieval

4.1 Introduction

Different traditions of research on WM have built on different preferred paradigms. Much research over the last five decades has used memory for lists of sequentially presented (mainly verbal) stimuli (for a review see Lewandowsky & Farrell, 2008). In recent years there was a surge of interest in WM for simultaneous displays of visual-spatial information (e.g. Luck & Vogel, 1997). The goal of the present study is to investigate commonalities between the two paradigms. In particular, we are interested in the role of contextual distinctiveness. In serial recall, temporal distinctiveness between successive list items plays a major role for memory performance. Here we ask whether spatial distinctiveness plays an analogous role in simultaneous-display paradigms.

4.1.1 Distinctiveness in Time and Space

Serial-order tasks are frequently used to investigate short-term or WM phenomena (Farrell & Lewandowsky, 2004; Henson, Norris, Page, & Baddeley, 1996; Nairne, 1988). In those tasks items are presented sequentially and have to be recalled in the order of presentation. Among the most frequent errors are transpositions, that is, recall of items in the wrong list positions. Transpositions follow the *locality constraint*: Items are more likely to be displaced to locations close to their original position (in particular neighboring positions) than more distant positions (Henson, et al., 1996; Lewandowsky & Farrell, 2008; Murdock & vom Saal, 1967).

The locality constraint can be explained by limited distinctiveness of WM representations on the time dimension. The most successful explanation for the locality constraint builds on the assumption that items are associated to position markers (i.e., representations of their list positions), and the position markers overlap to the extent that they are close in time (Burgess & Hitch, 1999; Lewandowsky & Farrell, 2008). Due to these

overlapping position markers, items are not only bound to the marker of their actual position but also to some extent to markers of neighboring positions. Hence, when using list position as retrieval cue to retrieve the corresponding item, the position markers also cue incorrect items to the degree that their positions overlap with the current position. As a consequence, people frequently transpose items with other items, and the proportion of such transpositions falls off with positional distance. The decline of transposition probability with distance is referred to as the *transposition gradient*.

In a typical visual WM task (e.g., Luck & Vogel, 1997; Zhang & Luck 2008), several visual objects are displayed simultaneously in different locations. Participants are required to memorize objects together with their spatial location to later compare them to a probe display, or to recall the object associated with a certain position in space. Hence, in these tasks, spatial location often serves as a retrieval cue, thus motivating the question whether there are transpositions in space similar to the temporal transpositions in serial-order tasks. Bays, Catalao, and Husain (2009) showed that confusions between objects explain a substantial proportion of errors in recall of items from simultaneous displays, but they do not assume a spatial gradient of distinctiveness between them. Our question is whether there is a locality constraint for confusions on the spatial dimension, equivalent to the locality constraint on the time dimension (Henson et al., 1996).

4.1.2 Ordinal or Metric Distinctiveness?

In the serial-order literature there is a debate on whether the temporal context is metric or ordinal (for a review see Morin, Brown, & Lewandowsky, 2010): The event-based approach favors the assumption that the temporal context advances only when a new event (e.g., a new list item) is encoded into WM, such that successive positions only reflect the order of events (Farrell & Lewandowsky, 2008; Lewandowsky & Brown, 2005). In contrast, the time-based approach assumes that the context advances as a function of time, such that the

relative lengths of the time gaps between successive items is reflected in different degrees of temporal distinctiveness between them (Brown, Neath, & Chater, 2007). Morin et al. (2010) summarize evidence for the importance of metric temporal distinctiveness in most paradigms of serial-order memory, with forward serial recall as a possible exception (see also Farrell, Wise, & Lelièvre, 2011). By analogy, the spatial effects in visual WM could be topological, reflecting only neighborhood relations between items, or metric, reflecting their relative distance in space. Here we investigate whether metric spatial distance affects the distinctiveness between representations in WM over and above the topological relation of neighborhood.

4.2 Experiments 1A and 1B

We used a local-recognition paradigm to investigate the effect of spatial proximity on false alarms. After encoding a display of six color items, participants were cued to the location of one item (the target) that they were to compare to a probe. Intrusion probes were probes that mismatched the target item but matched one of the other five memory items; these probes were to be rejected. We investigated whether the rate of false alarms to intrusion probes increases with the proximity between the target and the item matching the intrusion probe. If representations in WM are spatially imprecise and overlapping, spatially close items should be more difficult to distinguish from the target than more distant items. Hence, intrusion probes matching close neighbors of the target should cause more interference than far neighbors, and far neighbors should cause more interference than non-neighbors. To examine whether this is the case, we manipulated the spatial distance between neighboring objects in the memory display.

4.2.1 Method

Participants

Thirty students (11 male) participated in Experiment 1A (mean age 25 years, range: 19–32). In Experiment 1B, thirty students (3 male) participated (mean age 22 years, range: 18–35). Participants received financial compensation or course credit. Four participants of Experiment 1B were excluded due to bad performance (< 65% correct)¹.

Materials and Procedure

Experiments were programmed in MATLAB, using the Psychophysics Toolbox (Brainard, 1997).

Participants completed a local recognition task, requiring them to compare one object from a multi-stimuli memory display to a central probe. Each trial started with a memory display containing six colored circles on a grey background, arranged on an imaginary circle around the screen center. Of the six distances between neighboring stimuli, three were close (40° on the imaginary circle) and three were far (80°), randomly allocated to neighbor pairs. After the memory display had been presented for one second, the screen turned grey for 700 ms. Then a white arrow (i.e., the retro-cue) identified the relevant memory item by pointing to one of the six former stimulus locations. The cue was centrally presented for 100 ms. After a grey post-cue interval of 700 ms, the probe display was shown, containing one colored circle, which had to be compared to the cued item. The probe remained on screen until the participant pressed the left arrow key for a match, or the right arrow key for a mismatch. Performance feedback was provided visually for 500 ms and was followed by a grey inter-trial interval of 1.5 s (see Figure 1).

¹ Analyses including all participants showed the same data pattern as the analyses with subjects excluded.

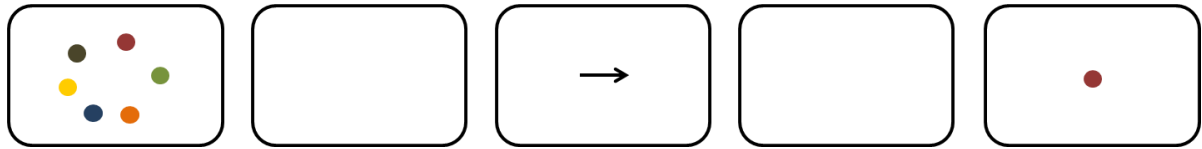


Figure 1. Description of a trial from Experiment 1A. The probe stimulus is a far neighbor in this example. The stimuli were presented on a grey background and the cues were presented in white. In Experiment 1B, white frames were presented around each memory stimulus and those frames remained on screen throughout each entire trial. Furthermore, the probe stimulus was surrounded by a white frame.

Experiment 1B was conducted to investigate whether the effect of spatial proximity observed in Experiment 1A was due to the inability to correctly identify the cued location. We replicated Experiment 1A with the only difference that white frames around each former stimulus location were presented throughout each trial, leaving no uncertainty about which location the cue pointed to and which item was the target for comparison.

Each participant completed 10 practice trials prior to the test phase. The test session consisted of 500 trials, split into 10 blocks. Fifty percent of trials were match trials. In mismatch trials, the probe was an intrusion probe (i.e., one of the not-cued memory items); either it was an item not neighboring the cued item (30% of trials) or it was a direct neighbor (20% of trials, with equal proportions of close and far neighbors)². Participants were instructed to respond as fast as possible without sacrificing accuracy.

4.2.2 Results

The average percentage correct was 84.5% (SD = 6.3) for Experiment 1A and 79.6% (SD = 4.3) for Experiment 1B. Trials associated with responses faster than 100 ms and slower than 7 s were excluded from RT analyses, as were error trials. RTs were log-transformed to reduce the skew of the distributions. Intrusion trials were sorted into non-neighbors, far

² Non-neighbors could have the same distance to the cued item as distant neighbors. However, this was very seldom the case; a non-neighbor with equal distance as a wide neighbor was probed in less than 3% of trials. Therefore, the amount of data points for non-neighbors with the same distance as distant neighbors was too little for statistical testing and no distinction between these different types of non-neighbor probes was drawn.

neighbors, and close neighbors. Match trials were omitted from analysis because they required a different response and are not of interest for the question investigated. When the sphericity assumption was violated, corrected Greenhouse-Geisser degrees of freedom (non-integer values) are presented.

A repeated-measures ANOVA on percentage correct revealed an effect of probe condition for Experiment 1A, $F(2, 58) = 15.342$, $p < .001$, partial $\eta^2 = .346$, and 1B, $F(2, 50) = 20.281$, $p < .001$, partial $\eta^2 = .448$. Planned contrasts showed the same picture for both experiments: Non-neighbor probes were rejected with higher accuracy than far neighbor probes, and far neighbors were rejected better than close neighbors (for statistics, see Table 1 and Figure 2).

Table 1

Planned contrasts on percentage correct for the different probe conditions for Experiments 1A and 1B

Contrast	F-Value	df1	df2	partial η^2	p
Experiment 1A					
non-neighbors vs. far neighbors	4.685	1	29	.139	.039
far neighbors vs. close neighbors	10.653	1	29	.269	.003
Experiment 1B					
non-neighbors vs. far neighbors	8.267	1	25	.248	.008
far neighbors vs. close neighbors	8.430	1	25	.252	.008

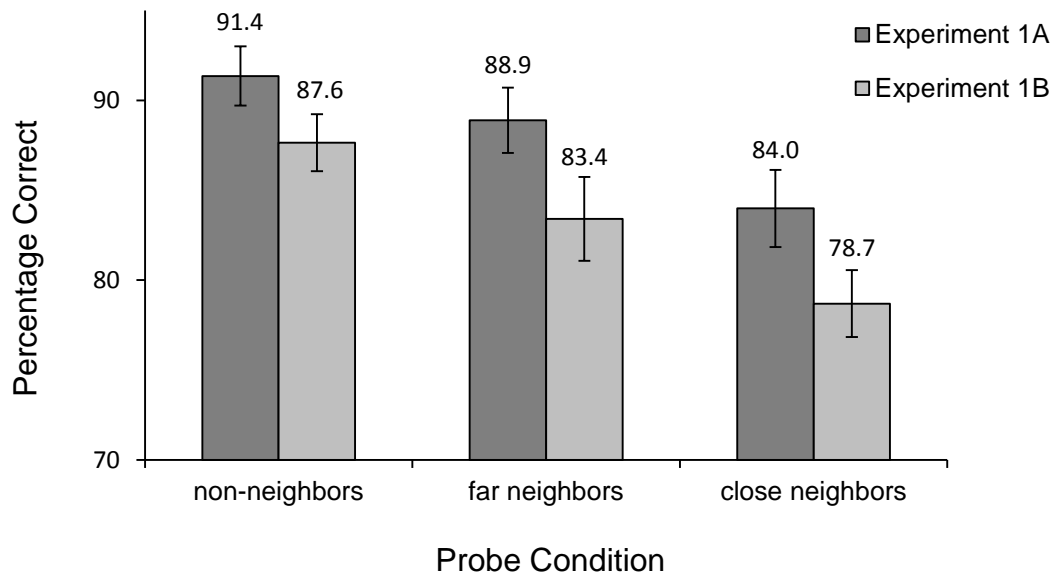


Figure 2. Accuracy data for the probe conditions in Experiment 1A and 1B. Error bars represent within-subject confidence intervals (95%).

The corresponding ANOVA for log-transformed RTs showed a significant effect of probe condition, $F(2, 58) = 4.122$, $p = .021$, partial $\eta^2 = .124$, for Experiment 1A, and a non-significant effect for Experiment 1B, $F(1.620, 40.492) = 2.227$, $p = .130$, partial $\eta^2 = .082$. Only the comparison between far and close neighbors in Experiment 1A reached significance in planned contrasts (far: .666 s, SD = .177; close: .692 s, SD = .169) (see Table 2 for statistics and Table 3 for means).

Table 2

Planned contrasts on log-transformed RT data for the different probe conditions for Experiments 1A and 1B

Contrast	F-Value	df1	df2	partial η^2	p
Experiment 1A					
non-neighbors vs. far neighbors	.008	1	29	.000	.930
far neighbors vs. close neighbors	6.474	1	29	.183	.017
Experiment 1B					
non-neighbors vs. far neighbors	.090	1	25	.004	.766
far neighbors vs. close neighbors	2.570	1	25	.093	.121

Table 3

Mean RTs for Experiment 1A and 1B

	Non-neighbors	Neighbors (far)	Neighbors (close)
Experiment 1A	.666 (.169)	.666 (.177)	.692 (.169)
Experiment 1B	.777 (.257)	.770 (.235)	.808 (.274)

Note. Standard deviations are provided in parentheses.

4.2.3 Discussion

The spatial closeness of the target item to the location of origin of an intrusion probe affects performance in a recognition task. This finding supports the assumption that spatial context representations in WM overlap as a function of their metric distance. Experiment 1B replicated the effect of spatial distance on accuracy observed in Experiment 1A, thereby ruling out the possibility that this effect was caused by an inability to correctly identify the cued location.

4.3 Experiment 2

In Experiment 2, we tested the analogy between spatial and temporal transposition errors more directly: We used a recall task, rendering Experiment 2 more similar to serial-recall experiments finding temporal transposition errors. Furthermore, we trace more precisely the recall probability as a function of Euclidian distance between the cued location and the recalled item. Participants saw a memory display with five colored stimuli, upon which one position was cued for recall. Participants selected the color of the cued item from an array of 12 colors. Two kinds of mistakes can be distinguished: transpositions and extra-list intrusions. Extra-list errors reflect the selection of a color not included in the memory display, whereas transpositions refer to the selection of another color from the memory display. We expect that transpositions are more frequent than extra-list errors, and that transpositions more frequently reflect the selection of items spatially close to the cued location than the selection of items further away.

In Experiment 2A, memory items were presented simultaneously, and in Experiment 2B they were presented sequentially. Sequential presentation served to test the possibility that spatial effects are caused by interference of spatially close items during encoding, rather than a spatial gradient of WM precision. For instance, perceptual crowding reduces the spatial resolution of individual items in dense arrays of multiple items, accounting for performance impairments in perception tasks (Cavanagh, 2004). If we find evidence for a spatial gradient with the sequential presentation, which removes any interference between items at encoding, the spatial distance effect cannot be attributed to encoding.

4.3.1 Method

Participants

Seventeen students (16 female, mean age = 23 years, range 20-30) participated in Experiment 2A, and 16 students (all female, mean age = 24 years; range: 21-30) in Experiment 2B.

Materials and Procedure

The memory display consisted of five colored squares placed in five cells of an invisible 9x9-grid on a grey background. The locations were selected at random with the constraint that summed horizontal and vertical distance between any two locations (i.e., their City-block distance) did not exceed 9 grid cells. The colors were selected at random from 12 colors (red, orange, yellow, brown, beige, olive, black, green, turquoise, blue, violet, pink), which were chosen to be easily distinguishable. Each color square was surrounded by a thin white frame. The only difference between Experiment 2A and 2B was the presentation of memory items: In Experiment 2A, all five memory items were presented simultaneously for one second, and in Experiment 2B, memory items were presented each for 500 ms seconds sequentially, so that only one item was visible on screen at any time. Following the presentation of the memory items, a grey screen with only the five white frames was shown for another second. After this retention interval, one of the frames was cued for recall by a thicker white frame for 100 ms, followed by a further one-second interval during which only the frames were visible. Then a palette of 12 colors was presented, arranged equidistantly on a virtual circle around the grid area. Participants had to select by a mouse click the color that had been presented in the cued location. Feedback was provided visually for 500 ms and followed by the instruction to press the space bar to continue. Another 500 ms after pressing the space bar the next trial began. Participants were instructed to answer as accurately as possible; speed was not emphasized. Figure 3 shows an example trial. Each participant completed two (Experiment

2A) or three (Experiment 2B) test sessions, including 24 practice trials prior to the first and 4 prior to the following session(s). The first test session comprised 400, the second 420 trials in Experiment 2A; in Experiment 2B all test session comprised 300 trials.

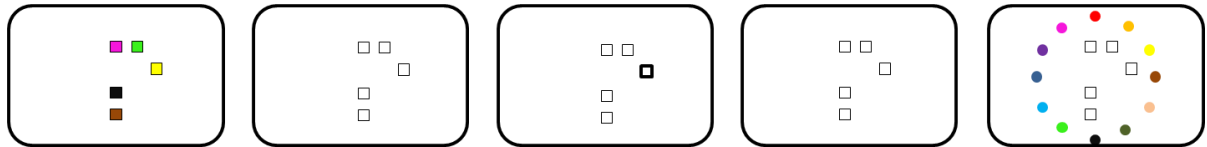


Figure 3. Flow of events of a trial from Experiment 2A. Stimuli were presented on a grey background and the frames as well as the cue (thicker frame) were presented in white. In Experiment 2B, stimuli were presented sequentially, one after the other with an inter-trial interval of 500 ms. The consequence was that at any one time, only one stimulus as visible in the memory display. Previously shown stimuli within the same trial were only designated by their remaining white frames.

4.3.2 Results

On average, participants in Experiment 2A committed a transposition error on 17.6% (SD = 8.3), and an extra-list intrusion on 13.0% of trials (SD = 5.5). In Experiment 2B, transposition errors occurred in 23.6% (SD = 6.4) and extra-list intrusions in 10.3% (SD = 4.8) of trials. Correctly answered trials were omitted from analyses. Analyses were conducted separately for both experiments.

The color palette contained 12 colors, one corresponding to the correct response, four to possible transpositions, and seven to possible extra-list errors. This implies a greater chance probability for selecting an extra-list item than a transposition. Therefore, we divided the proportion of each error category by the number of possibilities for that error category (i.e., 4 for transpositions, and 7 for extra-list intrusions). The corrected proportions show a clear preponderance of transposition errors (mean = 4.39, SD = 2.08) over extra-list errors (mean = 1.85, SD = 0.78) in Experiment 2A. This difference was significant, $t(16) = 5.058$, $p < .001$. Experiment 2B replicated this pattern, revealing significantly more transposition errors (mean = 5.89, SD = 1.59) than extra-list errors (mean = 1.47, SD = 0.69), $t(15) = 12.802$, $p < .001$.

To analyze whether the relative spatial distance between the cued location and the selected item affects performance, we calculated the Euclidian and the City-block distance between the cued item and the selected item for all transposition errors, resulting in 27 transposition distances for the Euclidian metric and 9 for the City-block metric.

Because on each trial the locations of the five memory items were chosen at random, the chances of errors in each distance class were not equal. To assess how frequently errors in each distance class, together with the class of extra-list errors, would be expected by chance, we ran a simulation for each participant's error trials, choosing among the 11 possible errors (i.e., the 12 colors except the correct one) with equal probability. Repeating the simulation 5,000 times generated a distribution of proportions of errors in each class of errors, which reflects the expected distribution for each class of errors under the null hypothesis that, on error trials, selections occur at random. To correct the observed data for chance, we divided the observed proportions of errors in each class by the corresponding mean proportion from the simulations, resulting in the corrected proportion of errors per class (Figure 4A for Experiment 2A and Figure 4B for Experiment 2B, solid line). The corrected data enabled us to examine whether the spatial distance between representations affects the distribution of transposition errors.

We tested three linear mixed-effect regression models to predict the corrected proportions of transposition errors by their transposition distance. The models were calculated with R (R development core team, 2007), using the lme4 package (Bates, 2005; Bates & Sarkar, 2007). Model 1 included only Euclidean distance as predictor, Model 2 included only City-block distance, and Model 3 included both predictors. Each model had subjects as a random factor on the intercept only, allowing for individual differences in overall performance. Model versions with additional random effects on the slopes, allowing individual differences also in the size of the distance effects, consistently resulted in a worse

fit than the corresponding model with random effects on the intercept only, according to the Bayesian Information Criterion (BIC).

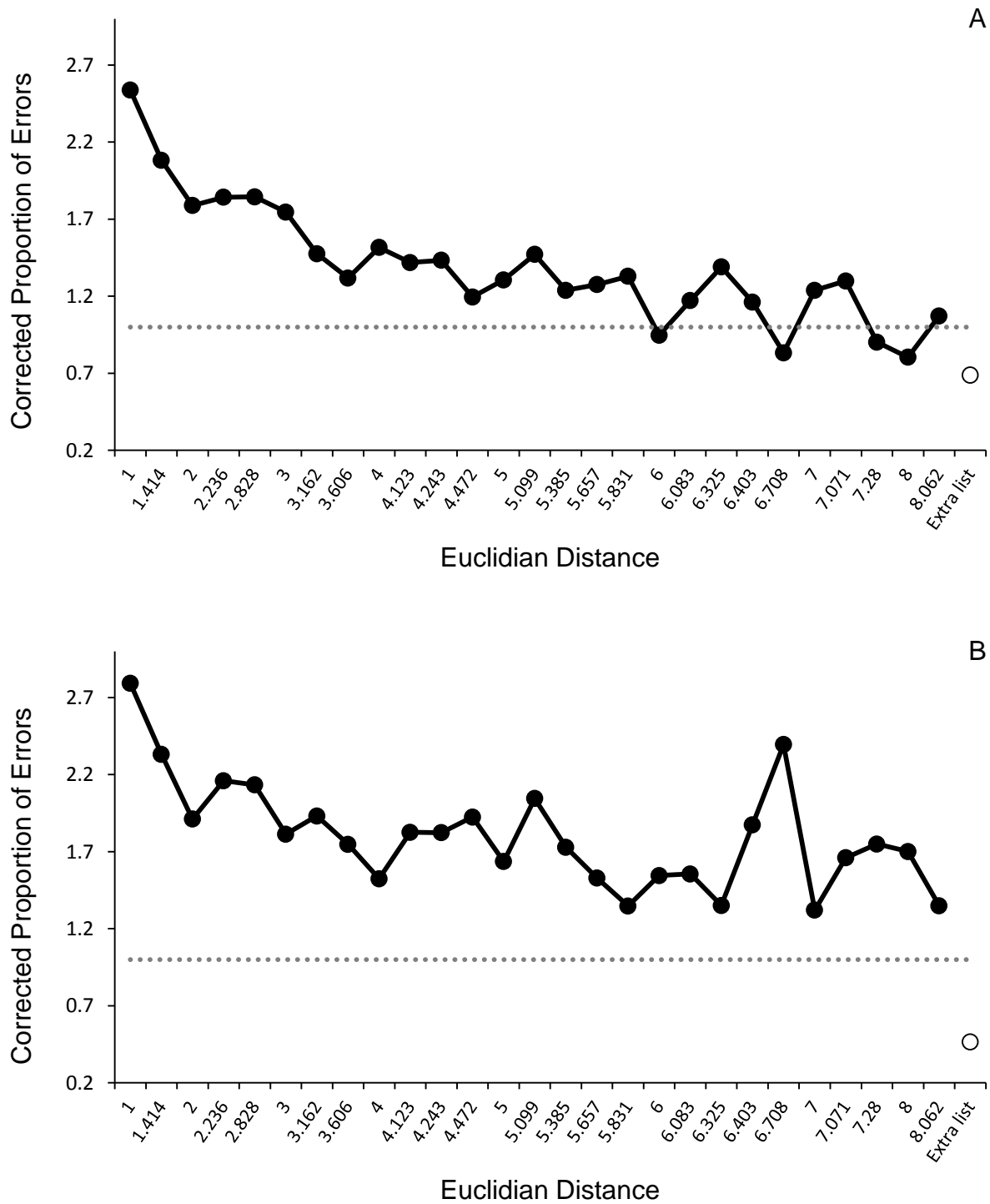


Figure 4. The x-axis shows the 27 Euclidian distance categories for transposition errors increasing from left to right. Extra-list errors are shown separately. The graphs show the data for Experiment 2A and panel B for Experiment 2B and display the proportion of errors corrected for chance as a function of spatial distances of transpositions and for extra-list errors. The grey dashed line represents performance predicted by chance. Panel A shows the data for Experiment 2A and panel B for Experiment 2B.

The best fit, as reflected by the smallest BIC value, was obtained for the model relying on Euclidian distance only (see Table 4). We calculated the p-values for the predictor (Euclidian distance) of this model, separately for Experiments 2A and 2B. In both cases, the p-values were 0.001 for the fixed factor Euclidian distance, indicating that Euclidean distance accounts for a significant share of the variance in corrected error proportions. Extra-list errors occurred significantly less than predicted by chance, which was confirmed by a significant t-test, comparing the corrected proportion of extra-list errors to chance performance (1), $t(16) = 6.041$, $p < .001$ for Experiment 2A and $t(15) = 17.482$, $p < .001$ for Experiment 2B.

Table 4

BIC and AIC values for the six models examined for Experiment 2A and 2B

	Fixed Factors	Random Parameters	BIC	AIC
Model 1	Euclidian	Participants	1280 / 1205	1263 / 1188
Model 2	Euclidian	Participants, Euclidian	1291 / 1212	1266 / 1188
Model 3	City Block	Participants	1287 / 1210	1270 / 1194
Model 4	City Block	Participants, City Block	1298 / 1220	1273 / 1196
Model 5	City Block, Euclidian	Participants	1290 / 1214	1269 / 1194
Model 6	City Block, Euclidian	Participants, City Block, Euclidian	1320 / 1242	1278 / 1201

Note. The first value before the slash refers to Experiment 2A, the value after the slash to Experiment 2B.

4.3.3 Discussion

Our results show that when items in WM are cued for recall by their spatial location, they are more likely to be transposed with close than with far neighbors, as shown by the spatial transposition gradients in Figure 4. The spatial gradient was obtained for Experiment 2A and 2B, thereby ruling out the explanation that the gradient was caused by interference

during encoding. The spatial gradient was better described as a decline of transposition likelihood with increasing Euclidean distance, rather than City-block distance. In addition, transposition errors occurred more often, and extra-list errors occurred less often, than would be expected by random guessing on error trials, consistent with the observation of Bays et al. (2009) that a substantial proportion of errors in visual WM tasks reflects retrieval of the wrong memory item, rather than random guessing.

4.4 General Discussion

Cue-based retrieval is among the best-established principles in memory research (Surprenant & Neath, 2009) and our data show that visual WM is no exception. Similarity between retrieval cues is a major cause of erroneous retrieval because competing retrieval candidates are less distinctive if they are associated to overlapping cues (Brown, et al., 2007). When memory items are displayed in different locations, and location cues are used to identify the item to be retrieved, then spatial location is the context cue for retrieval. As a consequence, spatial proximity plays the role of cue overlap, and as such determines the likelihood of confusions between memory contents.

More specifically, our results show that metric distance, not just ordinal distance, determines the likelihood of confusions between items bound to different locations. This finding is analogous to the finding that temporal distance, not just ordinal distance, affects retrieval accuracy in many, though not all, paradigms testing memory for serial order (Morin et al., 2010).

In sum, retrieval for recall and for local recognition is driven by contextual cues, and these cues overlap, so that they cue, to some extent, other items than the correct one. Depending on which dimension serves to discriminate the cues for the target item from cues to other, competing items, the relevant overlap can occur in time, in space, or in any other dimension. As a consequence, transpositions along the dimension that discriminates retrieval

cues are an important source of error in WM. This source so far plays no role in an influential model of visual WM, the slot model (Luck & Vogel, 1997; Zhang & Luck, 2008, 2011). This model assumes that each representation of an item is either stored in one of a limited number of discrete slots, or else is completely lost to WM. The slot model has frequently been applied to paradigms testing memory for a single item identified by its spatial position (Zhang & Luck, 2008, 2011). In these applications, all responses not reflecting memory for the item in the probed position have been regarded as random guesses. Our results show that when items are probed by their spatial location, a substantial number of spatial transposition errors occur. Transposition errors result in responses that do not reflect information about the probed item but instead reflect information about another item of the memory display (e.g., people report the color of a neighbor of the probed item, rather than the color of the probed item). These responses would be mis-attributed to random guessing by the slot model. Errors reflecting information about other items than the probed one have already been reported by Bays, Catalao, and Husain (2009). Those authors have not identified the source of this kind of errors. Here we determined that erroneous reports reflecting non-probed memory items arise systematically from the spatial imprecision of memory for each item's location. As a consequence of mis-attributing such errors to random guessing, applications of the slot model to paradigms testing memory for items in specific locations underestimate the capacity of visual WM.

One way to augment the slot model to incorporate transposition errors is to assume that every item stored in a slot is stored with a certain degree of imprecision not only for the feature to be recalled (e.g., color), but also for the feature that serves as a retrieval cue (e.g., spatial location). Retrieval of the content of a slot could then be described as a two-stage retrieval process (cf. Henson, 1998). In the first stage, the spatial location cue would be matched against the spatial features of the items stored in every slot. The slot with the highest

match would be selected for read-out. In the second stage, the to-be retrieved feature (e.g., color) would be reported from the selected slot according to that feature's precision. Transposition errors occur in the first stage when, due to the imprecision of spatial information in the slots, the wrong slot is selected for read-out.

5 References

- Bao, M., Li, Z. H., Chen, X. C., & Zhang, D. R. (2006). Backward inhibition in a task of switching attention within verbal working memory. *Brain Research Bulletin*, 69, 214-221.
- Bates, D. M. (2005). Fitting linear mixed models in R. *R News*, 5, 27–30.
- Bates, D. M., & Sarkar, D. (2007). *lme4: Linear mixed-effects models using S4 classes*, R package version 0.99875-6.
- Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, 321, 851-854. doi:10.1126/science.1158023
- Bays, P. M., Catalao, R. F. G., Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, 9, 1-11.
- Bays, P. M., Gorgoraptis, N., Wee, N., Marshall, L., & Husain, M. (2011). Temporal dynamics of encoding, storage, and reallocation of visual working memory. *Journal of Vision*, 11, 1-15. doi:10.1167/11.10.6
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433-436.
- Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review*, 114, 539-576.
- Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review*, 107, 127-181.
- Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review*, 106, 551-581.
- Cavanagh, P. (2004). Attention routines and the architecture of selection. In M. I. Posner (Ed.), *Cognitive neuroscience of attention* (pp.13-28). New York, NY: Guilford Press.

- Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin*, 104, 163-191.
- Cowan, N. (1995). *Attention and memory: An integrated framework*. New York: Oxford University Press.
- Cowan, N. (1999). An embedded-process model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). New York, NY: Cambridge University Press.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87–185.
- Cowan, N. (2005). *Working memory capacity*. New York, NY, USA: Psychological Press.
- Cowan, N. (2010). The magical mystery four: how is working memory capacity limited, and why? *Current Directions in Psychological Science*, 19, 51-57.
doi:10.1177/0963721409359277
- Cowan, N., Rouder, J. N., Blume, C. L., & Saults, J. S. (2012). Models of verbal working memory capacity: What does it take to make them work? *Psychological Review*, 119, 480-499. doi:10.1037/a0027791
- Cowan, N., Saults, J. S., Elliott, E. M., & Moreno, M. V. (2002). Deconfounding serial recall. *Journal of Memory and Language*, 46, 153-177.
- Delvenne, J.-F., Cleeremans, A., Laloyaux, C. (2010). Feature bindings are maintained in visual short-term memory without sustained focused attention. *Experimental Psychology*, 57, 108-116. doi:10.1027/1618-3169/a000014.
- Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review*, 9, 59-79.

- Farrell, S., & Lewandowsky, S. (2004). Modelling transposition latencies: Constraints for theories of serial order memory. *Journal of Memory and Language*, 51, 115-135.
- Farrell, S., Wise, V., & Lelièvre, A. (2011). Relations between timing, position, and grouping in short-term memory. *Memory & Cognition*, 39, 573-587.
- Garavan, H. (1998). Serial attention within working memory. *Memory & Cognition*, 26, 263-276.
- Glover, S., & Dixon, P. (2004). Likelihood ratios: A simple and flexible statistic for empirical psychologists. *Psychonomic Bulletin & Review*, 11, 791-806.
- Göthe, K., & Oberauer, K. (2008). The integration of familiarity and recollection information in short-term recognition: Modeling speed-accuracy trade-off functions. *Psychological Research*, 72, 289-303.
- Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations. *Journal of Cognitive Neuroscience*, 15, 1176-1194.
- Henson, R. N. A. (1998). Item repetition in short-term memory: Ranschburg repeated. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 1162-1181.
- Henson, R. N. A. (1998). Short-term memory for serial order: The start-end model. *Cognitive Psychology*, 36, 73-137.
- Henson, R. N. A., Norris, D. G., Page, M. P. A., & Baddeley, A. D. (1996). Unchained memory: Error patterns rule out chaining models of immediate serial recall. *Quarterly Journal of Experimental Psychology*, 49A, 80-115.
- Hübner, R., & Druey, M. D. (2006). Response execution, selection, or activation: What is sufficient for response-related repetition effects under task shifting? *Psychological Research*, 70, 245-261. doi:10.1007/s00426-005-0219-8

- Johnson, M. K. (1992). MEM: Mechanisms of recollection. *Journal of Cognitive Neuroscience*, 4, 268-280.
- Johnson, M. K., & Hirst, W. (1993). MEM: Memory subsystems as processes. In A. F. Collins, S. E. Gathercole, M. A. Conway, & P.E. Morris (Eds.), *Theories of Memory*, (pp. 241-286). East Sussex, England: Lawrence Erlbaum Associates.
- Johnson, M. K., Raye, C. L., Mitchell, K. J., Greene, E. J., Cunningham, W. A., & Sanislow, C. A. (2005). Using fMRI to investigate a component process of reflection: Prefrontal correlates of refreshing a just-activated representation. *Cognitive, Affective, & Behavioral Neuroscience*, 5, 339-361.
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual difference in working memory. *Psychological Review*, 99, 122-149.
- Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Science*, 4, 138-147.
doi:10.1016/S1364-6613(00)01452-2
- Koch, I., Gade, M., Philipp, A. M., & Schuch, S. (2010). The role of inhibition in task switching: A review. *Psychonomic Bulletin & Review*, 17, 1-14.
doi:10.3758/PBR.17.1.1
- Kuo, B. C., Rao, A., Lepsien, J., & Nobre, A. C. (2009). Searching for targets within the spatial layout of visual short-term memory. *The Journal of Neuroscience*, 29, 8032-8038. doi:10.1523/JNEUROSCI.0952-09.2009
- Kuo, B. C., Stokes, M., & Nobre, A. C. (2012). Attention modulates maintenance of representations in visual short-term memory. *Journal of Cognitive Neuroscience*, 24, 51-61. doi:10.1162/jocn_a_00087
- Landman, R., Spekreijse, H., & Lamme, A. F. (2003). Large capacity storage of integrated objects before change blindness. *Vision Research*, 43, 149-164.

- Lepsien, J., Griffin, I. C., Devlin, J. T., & Nobre, A. C. (2005). Directing spatial attention in mental representations: Interactions between attentional orienting and working memory load. *NeuroImage*, 26, 733-743.
- Lewandowsky, S., & Brown, G. D. A. (2005). Serial recall and presentation schedule: A micro-analysis of local distinctiveness. *Memory*, 13, 283-292.
- Lewandowsky, S., & Farrell, S. (2008). Short-term memory: New data and a model. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation* (Vol. 49, pp. 1-48). London, UK: Elsevier.
- Lewis-Peacock, J., Drysdale, A., Oberauer, K., & Postle, B. (2012). Neural evidence for a distinction between short-term memory and the focus of attention. *Journal of Cognitive Neuroscience*, 24, 61–79. doi:10.1162/jocn_a_00140
- Lin, P.-H., & Luck, S. J. (2012). Proactive interference does not meaningfully distort visual working memory capacity estimates in the canonical change detection task. *Frontiers in Psychology*, 3, 1-9. doi:10.3389/fpsyg.2012.00042
- Lovett, M., Reder, L. M., & Lebiere, C. (1999). Modeling working memory in a unified architecture: An ACT-R perspective. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 135–182). New York, NY: Cambridge University Press.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279–281.
- Makovski, T. (2012). Are multiple visual short-term memory storages necessary to explain the retro-cue effect? *Psychonomic Bulletin & Review*, 19, 470-476. doi:10.3758/s13423-012-0235-9
- Makovski, T., & Jiang, Y. V. (2007). Distributing versus focusing attention in visual short term memory. *Psychonomic Bulletin & Review*, 14, 1072-1078.

- Makovski, T., & Jiang, Y. V. (2008). Proactive interference from items previously stored in visual working memory. *Memory & Cognition*, *36*, 43-52.
- Makovski, T., Sussman, R., & Jiang, Y. V. (2008). Orienting attention in visual working memory reduces interference from memory probes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 369-380.
- Masson, M. E. J., (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavioral Research Methods*, *43*, 670-690. doi:10.3758/s13428-010-0049-5
- Matsukura, M., & Hollingworth, A. (2011). Does visual short-term memory have a high-capacity stage? *Psychonomic Bulletin & Review*, *18*, 1098–1104. doi:10.3758/s13423-011-0153-2
- Matsukura, M., Luck, S. J., & Vecera, S. P. (2007). Attention effects during visual short-term memory maintenance: protection or prioritization? *Perception & Psychophysics*, *69*, 1422-1434.
- Maxcey-Richard, A. M., & Hollingworth, A. (2012). The strategic retention of task-relevant objects in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. doi:10.1037/a0029496
- Mayr, U., & Keele, S. W. (2000). Changing internal constraints on action: the role of backward inhibition. *Journal of Experimental Psychology: General*, *129*, 4-26.
- McElree, B. (2001). Working memory and focal attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 817-835.
- McElree, B., & Doshier, B. A. (1989). Serial position and set size in short-term memory: The time course of recognition. *Journal of Experimental Psychology: General*, *118*, 346-373.

- Morey, C. C., & Cowan, N. (2004). When visual and verbal memories compete: Evidence of cross-domain limits in working memory. *Psychonomic Bulletin & Review*, *11*, 296-301.
- Morin, C., Brown, G. D. A., & Lewandowsky, S. (2010). Temporal isolation effect in recognition and serial recall. *Memory & Cognition*, *38*, 849-859.
- Murdock, B. B., & vom Saal, W. (1967). Transpositions in short-term memory. *Journal of Experimental Psychology*, *74*, 137-143.
- Nairne, J. S. (1988). A framework for interpreting recency effects in immediate serial recall. *Memory & Cognition*, *16*, 343-352.
- Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition*, *18*, 251-269.
- Nobre, A. C., Griffin, I. C., & Rao, A. (2008). Spatial attention can bias search in visual short term memory. *Frontiers in Human Neuroscience*, *1*, 1-9.
doi:10.3389/neuro.09.004.2007
- Oberauer, K. (2001). Removing irrelevant information from working memory: A cognitive aging study with the modified Sternberg task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 948-957.
- Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 411-421.
- Oberauer, K. (2003). Understanding serial position curves in short-term recognition and recall. *Journal of Memory and Language*, *49*, 469-483.
- Oberauer, K. (2005a). Binding and inhibition in working memory: Individual and age differences in short-term recognition. *Journal of Experimental Psychology: General*, *134*, 368-387.

- Oberauer, K. (2005b). Control of the contents of working memory - A comparison of two paradigms and two age groups. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 714-721.
- Oberauer, K. (2006). Is the focus of attention in working memory expanded through practice? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 197-214.
- Oberauer, K. (2008). How to say no: Single- and dual-process theories of short-term recognition tested on negative probes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 439-459.
- Oberauer, K. (2009). Design for a working memory. *Psychology of Learning and Motivation*, *51*, 45-100. doi:10.1016/S0079-7421(09)51002-X
- Oberauer, K., & Bialkova, S. (2009). Accessing information in working memory: Can the focus of attention grasp two elements at the same time? *Journal of Experimental Psychology: General*, *138*, 64-87.
- Oberauer, K., & Hein, L. (2012). Attention to information in working memory. *Current Directions in Psychological Science*, *21*, 164-169. doi:10.1177/0963721412444727
- Oberauer, K., & Lange, E. (2009). Activation and binding in verbal working memory: A dual-process model for the recognition of nonwords. *Cognitive Psychology*, *58*, 102-136.
- Oberauer, K., & Lewandowsky, S. (2008). Forgetting in immediate serial recall: Decay, temporal distinctiveness, or interference? *Psychological Review*, *115*, 544-576. doi:10.1037/0033-295X.115.3.544
- Oberauer, K., Lewandowsky, S., Farrell S., Jarrold, C., & Greaves, M. (2012). Modeling working memory: An interference model of complex span. *Psychonomic Bulletin & Review*, Advance online publication. doi:10.3758/s13423-012-0272-4

- Oberauer, K., Souza, A. S., Druey, M. D., & Gade, M. (2012). *Analogous mechanisms of selection and updating in declarative and procedural working memory: Experiments and a computational model*. Manuscript submitted for publication.
- Page, M. P. A., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review*, *105*, 761-781. doi:10.1037/0033-295X.105.4.761-781
- Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies, *Spatial Vision*, *10*, 437-442.
- Pratt, J., Kingstone, A., & Khoe, W. (1997). Inhibition of return in location- and identity based choice decision tasks. *Attention, Perception, & Psychophysics*, *59*, 964-971. doi:10.3758/BF03205511
- R development core team (2007). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing, <http://www.R-project.org>.
- Raye, C. L., Johnson, M. K., Mitchell, K. J., Greene, E. J., & Johnson M. R. (2007). Refreshing: A minimal executive function. *Cortex*, *43*, 135-145.
- Rotello, C. M., & Heit, E. (1999). Two-process models of recognition memory: Evidence for recall-to-reject? *Journal of Memory and Language*, *40*, 432-453.
- Schmidt, B. K., Vogel, E. K., Woodman, G. F., & Luck, S. J. (2002). Voluntary and automatic attentional control of visual working memory. *Perception & Psychophysics*, *64*, 754-763.
- Sligte, I. G., Scholte, H. S., & Lamme, V. A. F. (2008). Are there multiple visual short term memory stores? *PLoS ONE*, *3*, 1-9. doi:10.1371/journal.pone.0001699
- Souza, A. S., Hein, L., & Oberauer, K. (2012). *Attending to items and sets in working memory: Examining explanations of the retro-cue effect*. Unpublished manuscript, Department of Psychology, University of Zurich, Zurich, Switzerland.
- Surprenant, A. M., & Neath, I. (2009). *Principles of memory*. New York: Taylor & Francis.

- Williams, M., & Woodman, G. F. (2012). Directed forgetting and directed remembering in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. doi:10.1037/a0027389
- Woodman, G. F., & Vecera, M. (2011). The cost of accessing an object's feature stored in visual working memory. *Visual Cognition*, 19, 1-12. doi:10.1080/13506285.2012.521140
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, 46, 441-517.
- Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, 453, 233-236.
- Zhang, W., & Luck, S. J. (2011). The number and quality of representations in working memory. *Psychological Science*, 22, 1434-1441.
- Zhang, W., & Luck, S. L. (2009). Sudden death and gradual decay in visual working memory. *Psychological Science*, 20, 423-428.

Acknowledgements

First, I sincerely thank Klaus Oberauer, my supervisor, for his support, and for never losing his nerve. He has always been supportive, giving me the chance to proof myself and opened many opportunities to broaden my knowledge and to integrate into the scientific community.

Furthermore, I want to thank the Cognitive Psychology Team, in alphabetical order, Claudia von Bastian, Anke Benker, Michel Druet, Corinne Frey, Miriam Gade, Hsuan-Yu Lin, and Alessandra Souza for their support with whatever was needed, for downplaying all failures and boosting all successes on the way to this thesis.

I thank my dear friends, Birgit and Micha Siebold, and my sisters, Mieke and Frauke Hein, for bearing my endless monologues on science and working memory and for successfully distracting me to think about other things as well.

Special thanks are due to Alessandra Souza, for all her help and support in all respects, for sharing the office, for being a mentor and a friend and for ensuring my physical survival by joining many sushi sessions.

Furthermore, I thank Matthias Hartmann, Rafael Huber, Iliana Karipidis, Eszter Montvai, and Jeremias Wicki (in alphabetical order) for their assistance in the data collection.

Finally, I thank Matúš Rerko for everything.

CURRICULUM VITAE

Laura Hein

PERSONAL INFORMATION

Name	Laura Hein
Date of Birth	March 7 th 1985
Citizenship	German

ACADEMIC STUDIES

2010-2012	PhD studies in Cognitive Psychology, Cognitive Psychology Unit, University of Zürich, Switzerland; Supervisor: Prof. Dr. Klaus Oberauer
2007-2010	Graduate studies in Psychology, University of Freiburg, Switzerland Degree: Master of Science in Psychology
2004- 2007	Undergraduate studies in Psychology, University of Freiburg, Switzerland Degree: Bachelor of Science in Psychology

EMPLOYMENT

Since 2010	Research Assistant, Cognitive Psychology Unit, University of Zurich, Switzerland
------------	--
